STRUCTURAL REQUIREMENTS FOR ENZYMATIC EFFICIENCY IN COFACTOR-INDEPENDENT DECARBOXYLATION

BY

BIJOY J. DESAI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor John A. Gerlt, Chair
Professor James H. Morrissey
Professor Satish K. Nair
Assistant Professor Rutilio A. Fratti

Abstract

Orotidine 5'-monophosphate decarboxylase (OMPDC) is the last enzyme in the *de novo* pyrimidine biosynthetic pathway. It catalyzes the decarboxylation of orotidine 5'-monophosphate (OMP) to uridine 5'-monophosphate (UMP), without the requirement of a cofactor. It provides catalytic rate enhancements of $\sim 10^{17}$, compared to the uncatalyzed reaction, making it one of the most efficient enzymes known (1). OMPDC has thus been a paradigm to understand enzymatic catalysis. Yet, many aspects of its catalytic mechanism are not well understood. The reaction coordinate consists of a vinyl carbanion/carbene intermediate (2). To characterize the interaction of the intermediate with the enzyme active site, I performed site-directed mutagenesis of Ser 127 to L-glycerate residue (Chapter 2). This was accomplished by *in vitro* protein expression using L-glycerate charged tRNA. Examination of kinetic parameters of this mutant revealed that the backbone interactions accounts for 10^2 fold of the total 10^{17} fold rate enhancement (3).

OMPDC achieves its extraordinary rate enhancements by ground-state destabilization (GSD) and transition state destabilization (TSS) (4, 5). These strategies, enforced in the Michaelis complex, are accompanied by conformational changes, as characterized by the Apo (E_o) and inhibitor bound (E_c) crystal structures. A hallmark of these conformational changes is the selective formation of interactions in the E_c form compared to E_o, by residues remote from the active site of the enzyme. The role of these conformational changes in OMPDC catalysis has not been fully understood.

Using a combination of site-directed mutagenesis, substrate analogs, enzyme kinetic analysis and x-ray crystallographic analysis, research in Chapter 3 attempts to elucidate the mechanism for the formation of E_c from E_o. By comparing the second order rate constants (k_{cat}/K_M) for decarboxylation of OMP and a substrate analog devoid of 5'-phosphate group (EO) for mutants of residues that interact with the 5'-phosphate group (Q185A and R203A) with those of wild-type enzyme, I have determined intrinsic binding energy (IBE) contributions of these residues. Using a similar analysis with mutants of remote residues that selectively form hydrogen bonding (T159V, R203A and Y206F) and hydrophobic (V182A) interactions in E_c, I have determined the energetic contribution of these interactions to the IBE. Combining each mutant of the "remote" interactions with that of phosphate binding interactions, I have established the energetic independence of the "remote" interactions from 5'-phosphate binding in increasing catalytic efficiency of the enzyme. Taken together, the results of this research have lead to a structural and kinetic model that explains the roles of substrate binding and conformational change in OMPDC catalysis (6).

In Chapter 4, I investigate whether FrsA can perform cofactor independent decarboxylation of pyruvate. FrsA, a member of α/β hydrolase superfamily, was reported to catalyze cofactor independent decarboxylation of pyruvate, in a manner similar to OMPDC (7). Our studies show convincing evidence that frsA is not capable of catalyzing decarboxylation of pyruvate and the earlier report is incorrect (8).

References

(1) Radzicka, A.; Wolfenden, R. A proficient enzyme. *Science* **1995***, 267*, 90-93.

(2) Lee, J. K.; Houk, K. N. A proficient enzyme revisited: the predicted mechanism for orotidine monophosphate decarboxylase. *Science* **1997***, 276*, 942-945.

(3) Desai, B. J.; Goto, Y.; Cembran, A.; Fedorov, A. A.; Almo, S. C.; Gao, J.; Suga, H.; Gerlt, J. Investigating the role of a backbone to substrate hydrogen bond in omp decarboxylase using a site-specific amide to ester substitution. *Proc. Nat. Acad. Sci. U.S.A.* **2014**, 42, 15066-15077.

(4) Chan, K. K.; Wood, B. M.; Fedorov, A. A.; Fedorov, E. V.; Imker, H. J.; Amyes, T. L.; Richard, J. P.; Almo, S. C.; Gerlt, J. A. Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: evidence for substrate destabilization. *Biochemistry* **2009***, 48*, 5518-5531.

(5) Iiams, V.; Desai, B. J.; Fedorov, A. A.; Fedorov, E. V.; Almo, S. C.; Gerlt, J. A. Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: importance of residues in the orotate binding site. *Biochemistry* **2011***, 50*, 8497-8507.

(6) Desai, B. J.; Wood, B. M.; Fedorov, A. A.; Fedorov, E. V.; Goryanova, B.; Amyes, T. L.; Richard, J. P.; Almo, S. C.; Gerlt, J. A. Conformational changes in orotidine 5'-monophosphate decarboxylase: a structure-based explanation for how the 5'-phosphate group activates the enzyme. *Biochemistry* **2012***, 51*, 8665-8678.

(7) Lee, K. J.; Jeong, C. S.; An, Y. J.; Lee, H. J.; Park, S. J.; Seok, Y. J.; Kim, P.; Lee, J. H.; Lee, K. H.; Cha, S. S. FrsA functions as a cofactor-independent decarboxylase to control metabolic flux. *Nat. Chem. Biol.* **2011***, 7*, 434-436.

(8) Kellett, W. F.; Brunk, E.; Desai, B. J.; Fedorov, A. A.; Almo, S. C.; Gerlt, J. A.; Rothlisberger, U.; Richards, N. G. Computational, structural, and kinetic evidence that Vibrio vulnificus FrsA is not a cofactor-independent pyruvate decarboxylase. *Biochemistry* **2013***, 52*, 1842-1844.

Dedicated to My Beloved Mother

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to the committee chair, Professor John Gerlt, for his excellent guidance and advice. He taught me the importance of self-criticism in producing state-of-the-art science, and encouraged me strive for the best when my drive was waning.

I am grateful to Dr. B. Mckay Wood for providing an intellectually stimulating work environment, and for helping me through the initial learning curve on OMP decarboxylase work. I would also like to thank Dr. Tina Amyes, who developed the robust assay platform for OMP decarboxylase. The detailed work presented here would not have been possible without her contribution.

In Ranjani Murali, I found a companion with whom I could share our various interests and passions. She supported and encouraged me. Her company is a source of personal and intellectual growth. I am grateful to Shyamal Subramanyam, Ambika Nadkarni and Suchithra Prabhu for their wonderful friendship, and to Steven Alfano for his love and companionship. Finally, I am indebted to my family for their constant and unflinching love and support, without which none of this would have been possible.

TABLE OF CONTENTS

CHAPTER 1: GENERAL INTRODUCTION

IMPORTANCE OF OMPDC IN ENZYMOLOGY

Orotidine 5'-monophosphate decarboxylase is the last enzyme in the de-novo pyrimidine biosynthesis pathway. It catalyzes the conversion of orotidine 5'-monophosphate to uridine 5'-monophosphate, an essential precursor to pyrimidine ribonucleotides used for RNA synthesis as well as pyrimidine deoxyribonucleotides used for DNA synthesis. The product, UMP, acts as a feedback regulator for the *de novo* pyrimidine biosynthesis pathway (**Figure 1**). Due to its essential role in cellular physiology it is found in all domains of life (**Figure 2**). Its pivotal function in metabolism, OMPDC is an important target for anti-microbial drug discovery efforts.

Since its discovery, OMPDC has captured the interest of enzymologists due to its extraordinary catalytic prowess. Decarboxylation reactions are inherently difficult to perform as they usually form negatively charged intermediates that require stabilization. Most decarboxylases found in nature utilize either metal, pyridoxal phosphate or thiamine pyrophosphate as cofactors to stabilize the negative charge. In contrast, decarboxylation of OMP would proceed via a negatively charged vinyl anion intermediate, with no apparent means of internal stabilization. The rate of decarboxylation of uncatalyzed reaction was estimated by Beak et. al. and later by Lewis et. al., to be 5×10^{-14} and 3.4×10^{-15} respectively, making the estimated rate-enhancement afforded by OMPDC $\sim 10^{16} - 10^{17}$ fold (**Figure 3**) (1, 2). Until the

discovery of uroporphyrinogen III decarboxylase, OMPDC was the most efficient enzyme known. In spite of more than two decades of studies on OMPDC, the structural contributors to the 10^{17} fold rate enhancement have not been elucidated.

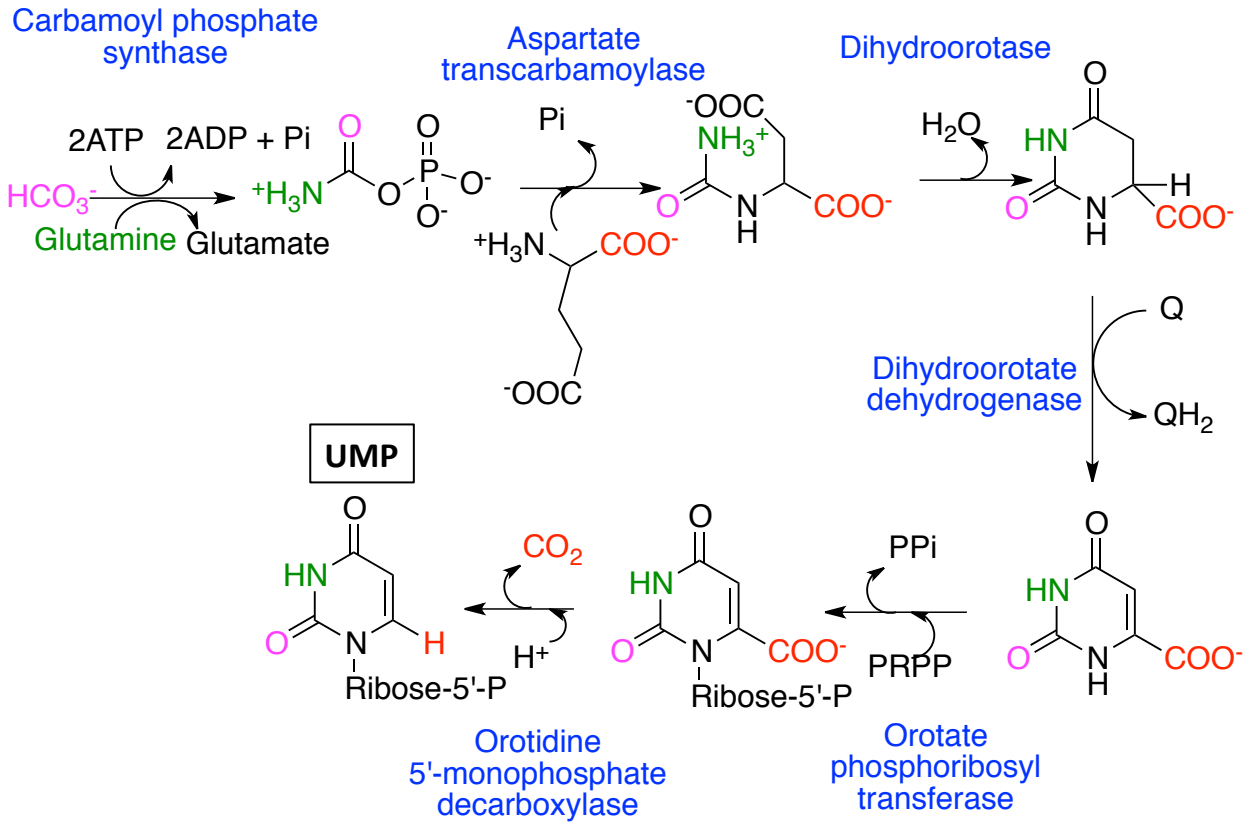


Figure 1. *De novo* pyrimidine biosynthesis pathway.

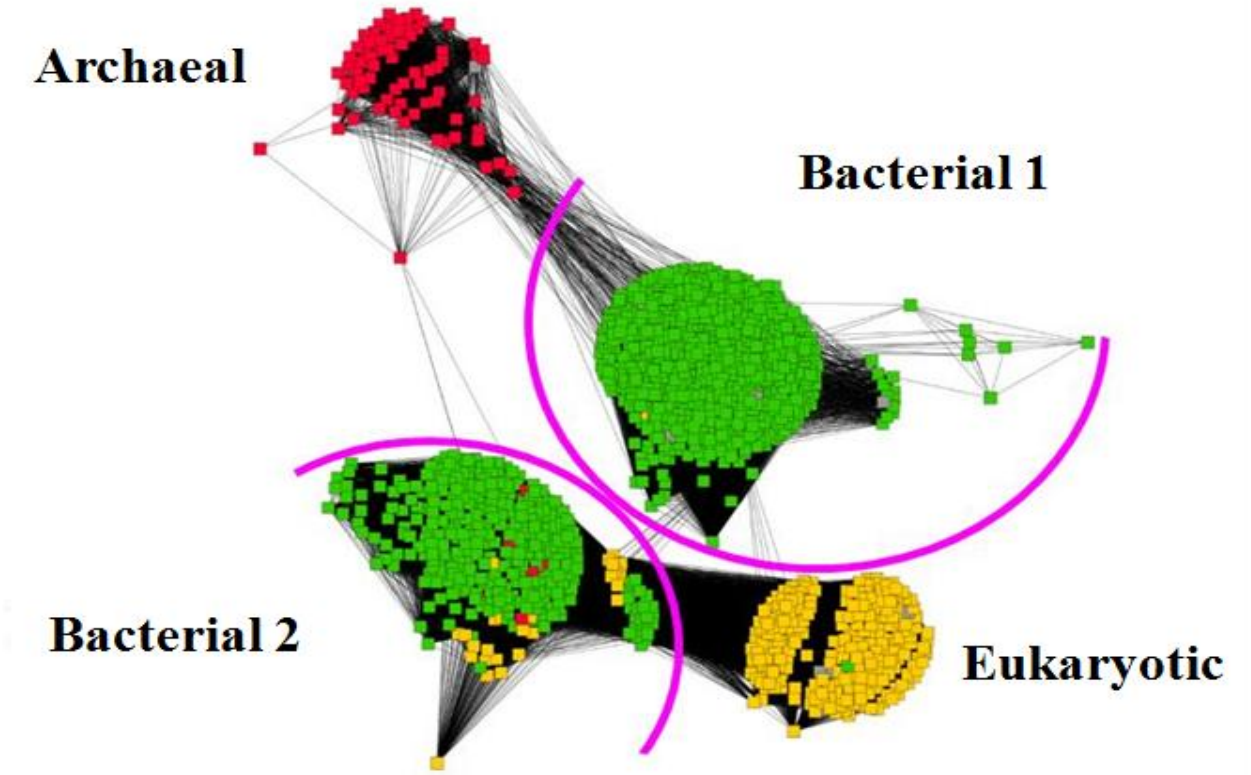


Figure 2. Sequence similarity network of all OMPDC protein sequences. Each node represents a unique protein sequence. Nodes that have sequences similarity less than or equal to a blast-e cutoff value of -80, are connected by a line. In this network archaeal (red), bacterial (green) and eukaryotic (yellow) OMPDCs separate into distinct clusters.

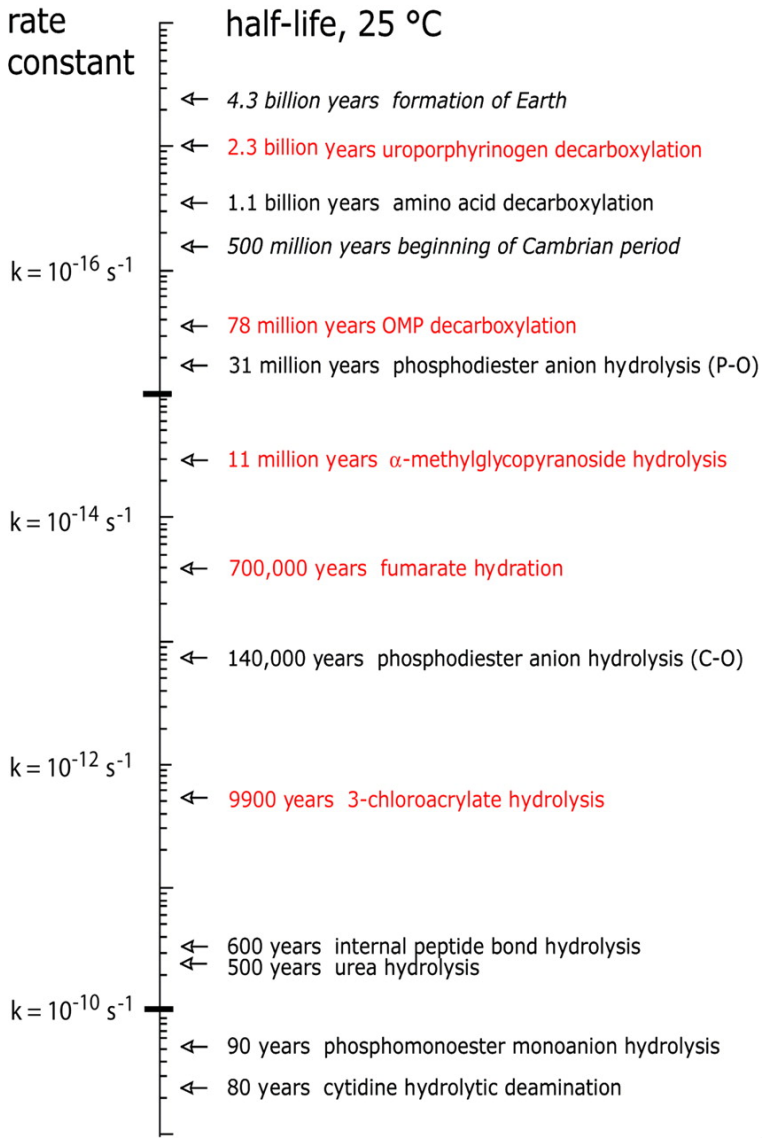


Figure 3. Half-lives of biochemically important reactions proceeding spontaneously in water. The figure was reproduced with the permission of Lewis *et al*. (3) Copyright belongs to the National Academy of Sciences of USA.

STRUCTURE OF OMPDC

Several research groups have independently solved crystal structures of OMPDC orthologs, from various organisms (4-6). All of the structures have divergent sequence identity, with key active site residues that are highly conserved. The model OMPDC enzyme used in this work is from the thermophilic archaeon, *Methanothermobacter*

thermoautotrophicum (MtOMPDC). This enzyme has high stability, can be purified in relatively large quantities and can easily crystallize to form well-diffracting crystals. This later property of this ortholog makes structural interpretation of kinetic properties of mutant OMPDCs easy. To date, there are more than 70 structures of MtOMPDC mutants in various inhibitor bound and apo forms.

OMPDC has a triose phosphate isomerase (TIM) $(\beta/\alpha)_8$-barrel fold, a common structural architecture found in enzymes (**Figure 4**). Examination of the inhibitor-bound structure shows that the enzyme is a homodimer, forming a two-fold axis of symmetry. Each monomer contains an active site located at the mouth of each $(\beta/\alpha)_8$-barrel. Each active site consists of residues donated by both symmetry related monomer, making OMPDC an obligate dimer.

The active site of OMPDC contains a highly conserved DXKXXD motif consisting of Asp70, Lys72 and Asp75*(*from symmetry related polypeptide) (**Figure 5**). This motif, along with Lys42 and Asp20, form a charged network. Asp70 has been implicated in ground-state destabilization (GSD) and Lys72 is the active site acid that protonates the vinyl carbanion intermediate after the departure of the carboxylate moiety (*vide infra*). Hence this motif is sometimes referred as the catalytic motif.

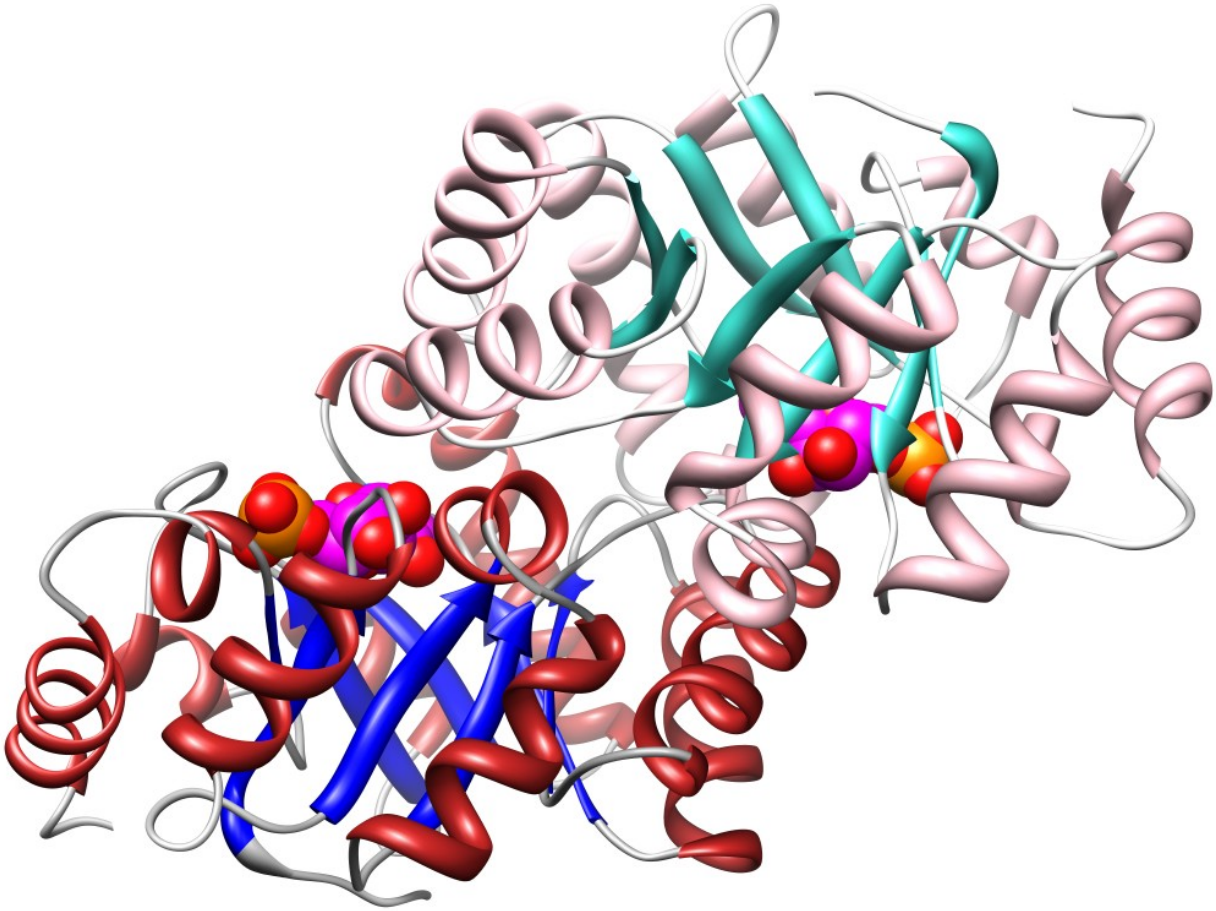


Figure 4. Structure of MtOMPDC bound to BMP (PDB ID: 3LTP). Protein is shown in cartoon representation and ligand BMP is shown in space-filling representation.

Apart from the catalytic motif, the active site of OMPDC contains several conserved interactions with the pyridine base and the phosphoribosyl moiety. Among these is Ser127 whose side-chain makes important contact with N-3 of the pyrimidine base and side-chain of Gln185; Gln185 also makes hydrogen-bonding interactions with O2 of the pyrimidine ring and the phosphate group of the phosphoribosyl moiety, "tethering", so to say, different parts of the enzyme and substrate together into a precise geometry

(**Figure 5**). Gln185 is a part of a flexible loop that is disordered in the apo structure and is ordered in the inhibitor/substrate bound structure, completely occluding the active site from solvent. Arg203 and Gly203 also make important contacts with the phosphate group. Apart form the residues that directly interact with the phosphate group; other residues in the active site indirectly interact with the phosphate group via ordered water molecules, forming an intricate network of interactions (not shown).

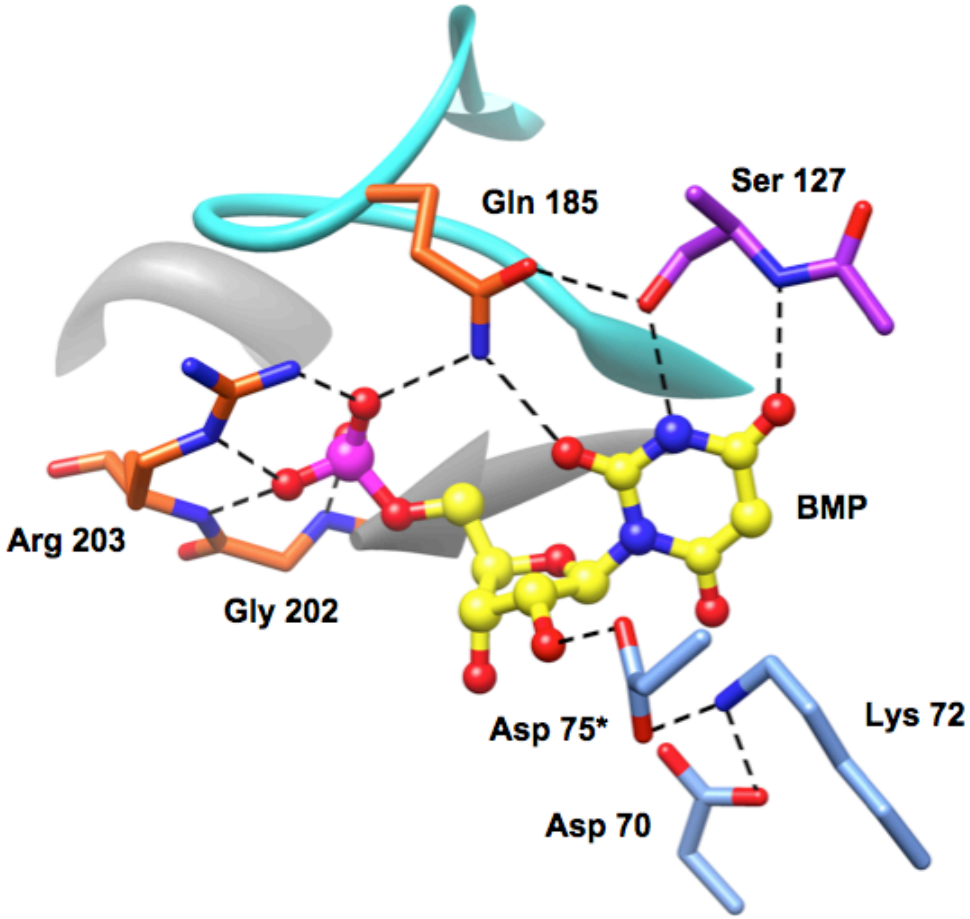


Figure 5. Important active site interactions in OMPDC. The figure was generated using PDB coordinates of MtOMPDC bound to BMP (PDB ID: 3LTP).

MECHANISM OF OMPDC

Prior to the availability of crystal structures, several mechanisms had been proposed for OMP decarboxylase catalyzed reaction (**Figure 6**); most of these mechanisms were incompatible with the solved structures of OMPDC due to absence of appropriate functional groups in the enzyme active site (1, 4, 7-10). The established mechanism consistent with the crystal structures and kinetic studies is a two-step mechanism consisting of formation of a vinyl carbanion intermediate and subsequent protonation (**Figure 7**). Earlier studies of comparing C-13 isotope effects of OMP vs 5-fluoro OMP (FOMP) by Van vleet *et al.* found ~1.5% decrease in isotope effect implying a stabilizing effect of carbanion intermediate by the inductive 5'-fluoro substituent of FOMP and thereby, indirectly showing importance of a carbanion intermediate (9). More direct evidence of the two-step carbanion mechanism came from work by Toth *et al.*, which showed that OMPDC catalyzed the H/D exchange (50%) of H6 of UMP in 50/50 H_2O/D_2O solution (10, 11).

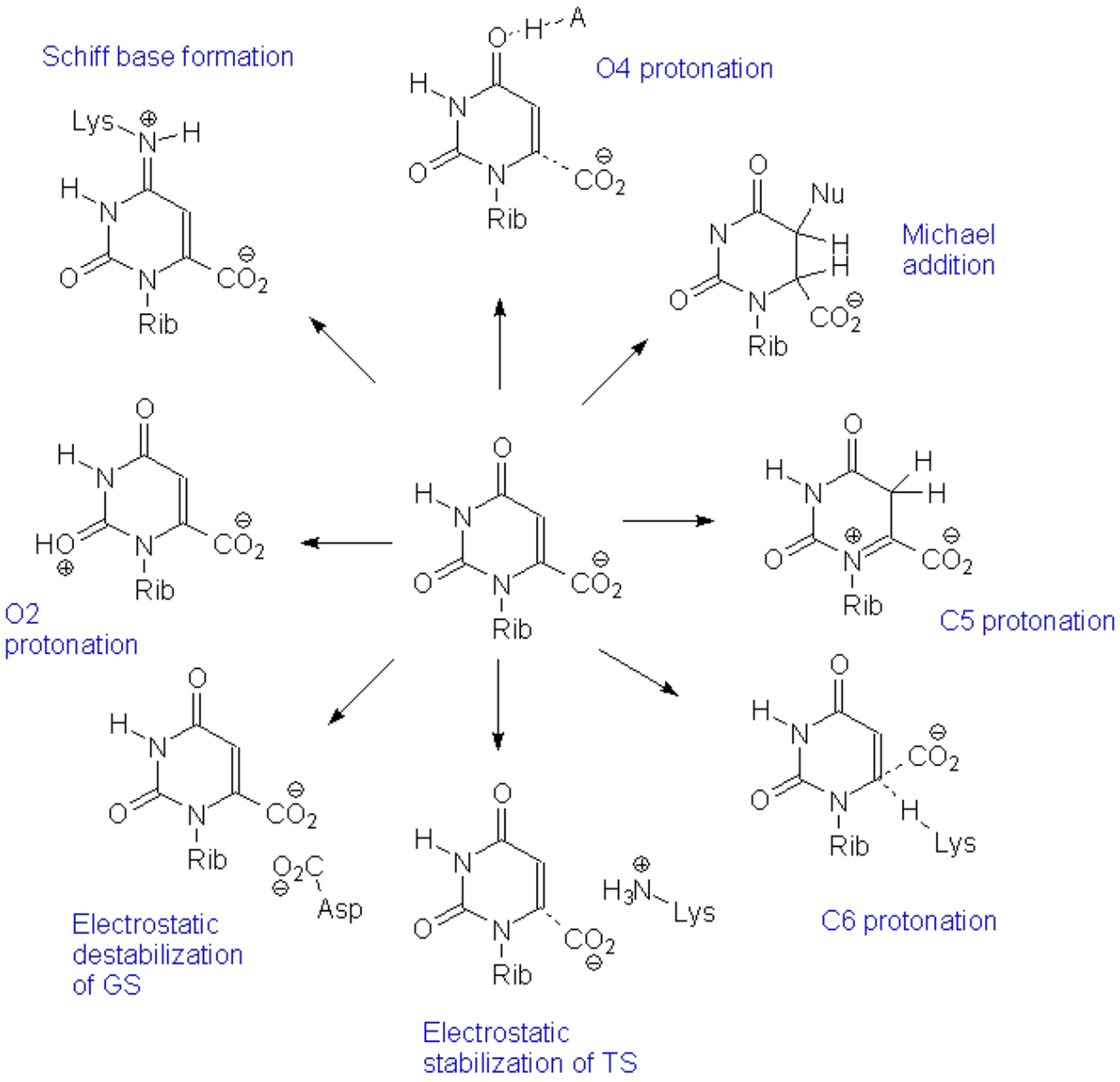


Figure 6. Proposed mechanisms for OMP decarboxylase reaction. This figure was reproduced with permission of Houk *et al. (8)* Copyright to this figure belongs to J. Phys. Chem. B.

Lee *et al.* proposed a carbene based transition state based on quantum mechanical calculations (7). This mechanism proposes a stabilization of the transition state by formation of resonance structure of vinyl carbanion where the negative charge resides

on O4 of the intermediate, which can then be stabilized by protonation at O4. Although the crystal structure of OMPDC (solved after the proposed mechanism) does not contain any proton donor near O4 of the pyrimidine base, the possibility of a carbene resonance form, stabilized by an enhanced hydrogen bond (short-strong hydrogen bond) with the backbone amide nitrogen of conserved active site Ser127, cannot be ruled out. Indirect support for the importance of such a resonance form comes form the discovery of the potent inhibitor 6-hydroxy UMP (BMP) (picomolar K_i), a putative transition-state analog for OMPDC catalyzed reaction (**Figure 7**) (12). Chapter 2 investigates this possibility by mutating of the backbone amide bond to ester bond and kinetically characterization of this mutant.

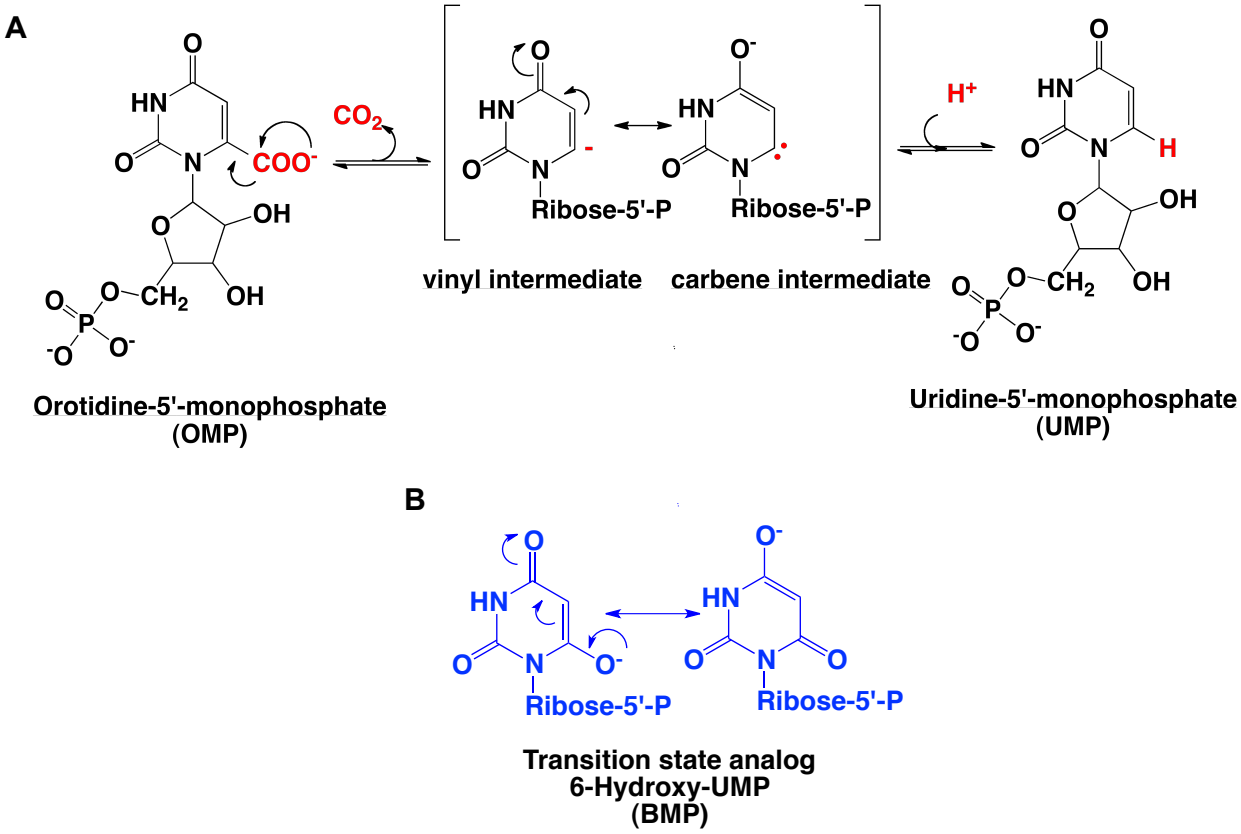


Figure 7. (A) Two-step mechanism of OMPDC catalyzed reaction showing vinyl and carbene intermediates. (B) Resonance forms of transition state analog BMP at pH 7.

ROLE OF GROUND-STATE DESTABILIZATION

After the first crystal structure of OMPDC bound to BMP was solved, it became apparent that OMP in the active site of the enzyme could experience electrostatic stress from the nearby negative charge of Asp70 (13). If this is true, ground-state destabilization of substrate carboxylate group can facilitate its departure and contribute to the 10^{17} fold rate enhancement. For some time, the existence of ground-state destabilization in OMPDC catalysis had been a subject of controversy. Molecular dynamic studies by Warshel *et al.* posit that only transition-state stabilization is important for OMPDC catalysis, as the charged preorganization of the active site reduces the protein-protein reorganization energy (14). In contrast, work done in our lab by Chan *et al.* showed that the D70N mutant has a reduced k_{cat} compared to wild-type enzyme for FOMP decarboxylation but has an unchanged rate of H/D exchange reaction of FUMP H-6 in deuterated solvent (15). This shows the possible role of Asp70 in ground state destabilization. Iams *et al.* showed the role of a hydrophobic pocket proximal to the orotate-binding site, formed by Ile96, Leu123, Val155, and Pro180, in accentuating the stress experienced by the carboxylate moiety (16). It was proposed that that OMP interacts with the enzyme using favorable binding energy of various active site interactions, remote from the carboxylate moiety. This juxtaposes the carboxylate moiety next to the unfavorable interactions of Asp70 and the hydrophobic pocket, facilitating its departure. This phenomenon, called the "Circe effect", where the binding energy of a remote substituent "pays" for unfavorable interaction at the site of chemistry, to facilitate catalysis, was first proposed by Jencks (17). More recent work by

Fujihashi *et al.* shows the presence of significant stress in crystal structure of MtOMPDC bound to various analogs of UMP with different C-6 substituent groups, supporting the presence of ground-state destabilization. They show that Lys72 is responsible for steric stress which is relieved upon decarboxylation (18, 19).

ROLE OF 5'-PHOSPHATE MOIETY OF OMP

The 5'-phosphate group of OMP is spatially remote from the carboxylate moiety. Yet it has been shown to be very important in OMPDC catalyzed reaction. The second order rate constants for a truncated substrate devoid of the 5'-phosphate group decreases by a factor of 2.5 X 10^8 (**Figure 8**) (20-23). This is the largest known effect of a remote substituent, not partaking in chemical transformation on enzyme-catalyzed rates. The 5'-phosphate moiety makes several contacts with active site residues including, but not limited to, side chains of Gln185 and Arg203 and backbone of Gly202. Together they form a network of interactions in the substrate/inhibitor-bound form of the enzyme. The importance of these interactions in OMPDC catalysis has not been well characterized and is one the main topics of inquiry in Chapter 3.

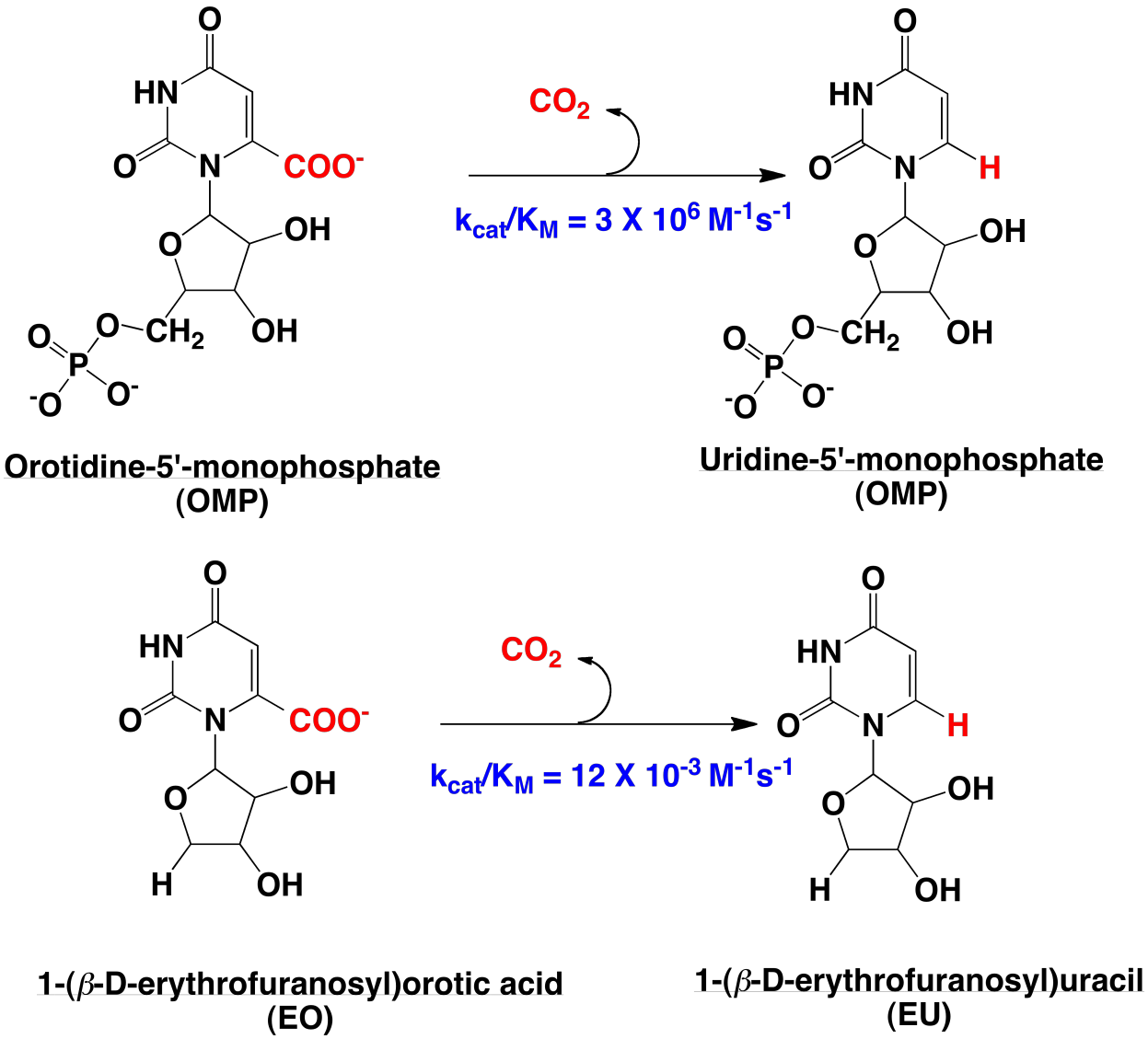


Figure 8. MtOMPDC catalyzed second order rates for decarboxylation of OMP and truncated substrate having no 5'-phosphate group (EO).

CONFORMATIONAL CHANGES IN OMPDC

Comparison of the apo form of the enzyme with the BMP bound form shows several changes in the organization of the enzyme. The most important is organization of the active site loop. Work in our lab by Wood *et al.* showed that ordering of the active site loop is facilitated by formation of a hydrophobic cluster consisting of Val182, Ile199,

Val201 and Ile218 (**Figure 9**). Alanine mutants of these hydrophobic residues show an increase in K_M, with little change in k_{cat}. It was proposed that these residues stabilize the closed conformation of the enzyme (24).

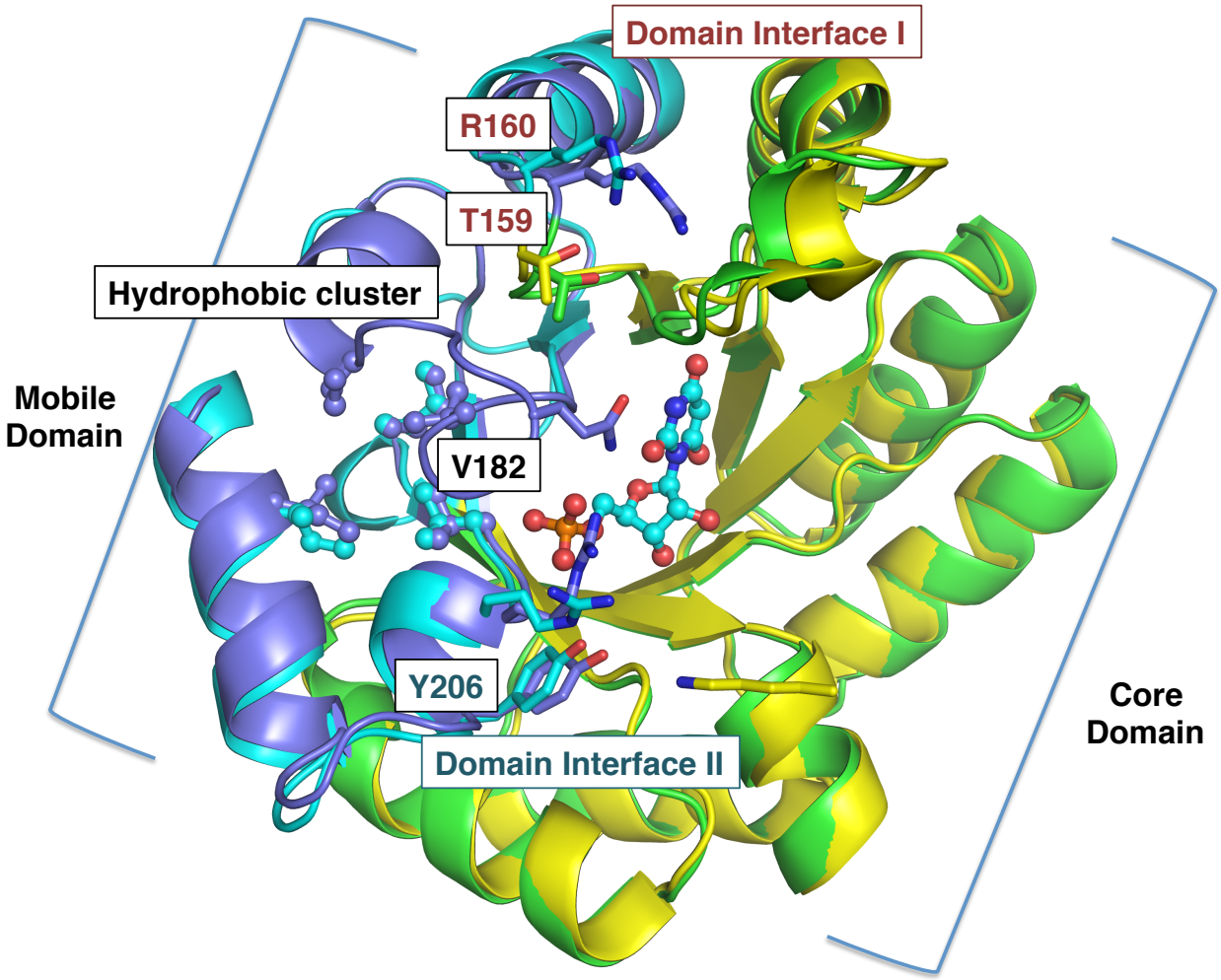


Figure 9. Superposition of MtOMPDC apo structure (PDB ID: 1LOR) with BMP bound structure (PDB ID: 3LTP). Only the monomer is shows for easy visualization.

Larsen et al., using HINGEFIND algorithm, classified the OMDPC architecture into two domains based on their flexibility: a) the rigid core domain consisting of two half $(\beta/\alpha)_8$-barrels associated with each other by the dimer interface; and b) the relatively flexible, mobile domain (25). Comparison of the apo and BMP bound structures of

MtOMPDC shows several hydrogen bonding interactions that occur at the interface of the two domains in the BMP bound structure (T159, R160 and Y206) (**Figure 9**). The detailed importance of these interactions is studied in Chapter 3. Moreover, the relationship between substrate binding via the phosphate moiety and remote conformational changes is also investigated.

Finally, the principles that are learned from OMPDC catalysis are applied in Chapter 4, where the possibility of FrsA, a protein belonging to α/β hydrolase superfamily to catalyze cofactor-less decarboxylation of pyruvate, is investigated. Prior to our work, Lee *et al.* reported that this protein decarboxylates pyruvate in a cofactor independent manner similar to OMPDC (26).

References

(1) Beak, P.; Siegel, B. Mechanism of decarboxylation of 1, 3-dimethylorotic acid. A model for orotidine 5'-phosphate decarboxylase. *J. Am. Chem. Soc.* **1976***, 98*, 3601-3606.

(2) Lewis, C. A.,Jr; Wolfenden, R. Orotic acid decarboxylation in water and nonpolar solvents: a potential role for desolvation in the action of OMP decarboxylase. *Biochemistry* **2009***, 48*, 8738-8745.

(3) Lewis, C. A.,Jr; Wolfenden, R. Uroporphyrinogen decarboxylation as a benchmark for the catalytic proficiency of enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **2008***, 105*, 17328-17333.

(4) Appleby, T. C.; Kinsland, C.; Begley, T. P.; Ealick, S. E. The crystal structure and mechanism of orotidine 5'-monophosphate decarboxylase. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 2005-2010.

(5) Houk, K. N.; Lee, J. K.; Tantillo, D. J.; Bahmanyar, S.; Hietbrink, B. N. Crystal structures of orotidine monophosphate decarboxylase: does the structure reveal the mechanism of nature's most proficient enzyme? *Chembiochem* **2001**, *2*, 113-118.

(6) Wu, N.; Pai, E. F. Crystal structures of inhibitor complexes reveal an alternate binding mode in orotidine-5'-monophosphate decarboxylase. *J. Biol. Chem.* **2002**, *277*, 28080-28087.

(7) Lee, J. K.; Houk, K. N. A proficient enzyme revisited: the predicted mechanism for orotidine monophosphate decarboxylase. *Science* **1997**, *276*, 942-945.

(8) Stanton, C. L.; Kuo, I. F.; Mundy, C. J.; Laino, T.; Houk, K. N. QM/MM metadynamics study of the direct decarboxylation mechanism for orotidine-5'-monophosphate decarboxylase using two different QM regions: acceleration too small to explain rate of enzyme catalysis. *J Phys Chem B* **2007**, *111*, 12573-12581.

(9) Van Vleet, J. L.; Reinhardt, L. A.; Miller, B. G.; Sievers, A.; Cleland, W. W. Carbon isotope effect study on orotidine 5'-monophosphate decarboxylase: support for an anionic intermediate. *Biochemistry* **2008**, *47*, 798-803.

(10) Toth, K.; Amyes, T. L.; Wood, B. M.; Chan, K.; Gerlt, J. A.; Richard, J. P. Product deuterium isotope effect for orotidine 5'-monophosphate decarboxylase:

evidence for the existence of a short-lived carbanion intermediate. *J. Am. Chem. Soc.* **2007**, *129*, 12946-12947.

(11) Toth, K.; Amyes, T. L.; Wood, B. M.; Chan, K.; Gerlt, J. A.; Richard, J. P. Product deuterium isotope effects for orotidine 5'-monophosphate decarboxylase: effect of changing substrate and enzyme structure on the partitioning of the vinyl carbanion reaction intermediate. *J. Am. Chem. Soc.* **2010**, *132*, 7018-7024.

(12) Levine, H. L.; Brody, R. S.; Westheimer, F. H. Inhibition of orotidine-5'-phosphate decarboxylase by 1-(5'-phospho-beta-d-ribofuranosyl)barbituric acid, 6-azauridine 5'-phosphate, and uridine 5'-phosphate. *Biochemistry* **1980**, *19*, 4993-4999.

(13) Wu, N.; Mo, Y.; Gao, J.; Pai, E. F. Electrostatic stress in catalysis: structure and mechanism of the enzyme orotidine monophosphate decarboxylase. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 2017-2022.

(14) Warshel, A.; Strajbl, M.; Villa, J.; Florian, J. Remarkable rate enhancement of orotidine 5'-monophosphate decarboxylase is due to transition-state stabilization rather than to ground-state destabilization. *Biochemistry* **2000**, *39*, 14728-14738.

(15) Chan, K. K.; Wood, B. M.; Fedorov, A. A.; Fedorov, E. V.; Imker, H. J.; Amyes, T. L.; Richard, J. P.; Almo, S. C.; Gerlt, J. A. Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: evidence for substrate destabilization. *Biochemistry* **2009**, *48*, 5518-5531.

(16) Iiams, V.; Desai, B. J.; Fedorov, A. A.; Fedorov, E. V.; Almo, S. C.; Gerlt, J. A. Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: importance of residues in the orotate binding site. *Biochemistry* **2011***, 50*, 8497-8507.

(17) Jencks, W. P. *Binding energy, specificity, and enzymic catalysis: the circe effect;* Wiley Online Library: .

(18) Fujihashi, M.; Mito, K.; Pai, E. F.; Miki, K. Atomic resolution structure of the orotidine 5'-monophosphate decarboxylase product complex combined with surface plasmon resonance analysis: implications for the catalytic mechanism. *J. Biol. Chem.* **2013***, 288*, 9011-9016.

(19) Fujihashi, M.; Ishida, T.; Kuroda, S.; Kotra, L. P.; Pai, E. F.; Miki, K. Substrate distortion contributes to the catalysis of orotidine 5'-monophosphate decarboxylase. *J. Am. Chem. Soc.* **2013***, 135*, 17432-17443.

(20) Amyes, T. L.; Richard, J. P.; Tait, J. J. Activation of orotidine 5'-monophosphate decarboxylase by phosphite dianion: the whole substrate is the sum of two parts. *J. Am. Chem. Soc.* **2005***, 127*, 15708-15709.

(21) Goryanova, B.; Amyes, T. L.; Gerlt, J. A.; Richard, J. P. OMP decarboxylase: phosphodianion binding energy is used to stabilize a vinyl carbanion intermediate. *J. Am. Chem. Soc.* **2011***, 133*, 6545-6548.

(22) Amyes, T. L.; Ming, S. A.; Goldman, L. M.; Wood, B. M.; Desai, B. J.; Gerlt, J. A.; Richard, J. P. Orotidine 5'-monophosphate decarboxylase: transition state

stabilization from remote protein-phosphodianion interactions. *Biochemistry* **2012***, 51*, 4630-4632.

(23) Goryanova, B.; Goldman, L. M.; Amyes, T. L.; Gerlt, J. A.; Richard, J. P. Role of a guanidinium cation-phosphodianion pair in stabilizing the vinyl carbanion intermediate of orotidine 5'-phosphate decarboxylase-catalyzed reactions. *Biochemistry* **2013***, 52*, 7500-7511.

(24) Wood, B. M.; Amyes, T. L.; Fedorov, A. A.; Fedorov, E. V.; Shabila, A.; Almo, S. C.; Richard, J. P.; Gerlt, J. A. Conformational changes in orotidine 5'-monophosphate decarboxylase: "remote" residues that stabilize the active conformation. *Biochemistry* **2010***, 49*, 3514-3516.

(25) Harris, P.; Poulsen, J. C.; Jensen, K. F.; Larsen, S. Substrate binding induces domain movements in orotidine 5'-monophosphate decarboxylase. *J. Mol. Biol.* **2002***, 318*, 1019-1029.

(26) Lee, K. J.; Jeong, C. S.; An, Y. J.; Lee, H. J.; Park, S. J.; Seok, Y. J.; Kim, P.; Lee, J. H.; Lee, K. H.; Cha, S. S. FrsA functions as a cofactor-independent decarboxylase to control metabolic flux. *Nat. Chem. Biol.* **2011***, 7*, 434-436.

CHAPTER 2: INVESTIGATING THE ROLE OF A BACKBONE TO SUBSTRATE HYDROGEN BOND IN OMP DECARBOXYLASE USING A SITE-SPECIFIC AMIDE TO ESTER SUBSTITUTION

ABSTRACT: Hydrogen bonds between backbone amide groups of enzymes and their substrates are often observed, but their importance in substrate binding and/or catalysis is not easy to investigate experimentally. We describe the generation and kinetic characterization of a backbone amide to ester substitution in the orotidine 5′-monophosphate (OMP) decarboxylase from *Methanobacter thermoautotrophicum* (MtOMPDC) to determine the importance of a backbone amide–substrate hydrogen bond. The MtOMPDC-catalyzed reaction is characterized by a rate enhancement (~10^{17}) that is among the largest for enzyme-catalyzed reactions. The reaction proceeds through a vinyl anion intermediate that may be stabilized by hydrogen bonding interaction between the backbone amide of a conserved active site serine residue (Ser-127) and oxygen (O4) of the pyrimidine moiety and/or electrostatic interactions with the conserved general acidic lysine (Lys-72). In vitro translation in conjunction with amber suppression using an orthogonal amber tRNA charged with L-glycerate (^{HO}S) was used to generate the ester backbone substitution (S127^{HO}S).

Reprinted (adapted) with permission of Desai, B. J.; Goto, Y.; Cembran, A.; Fedorov, A. A.; Almo, S. C.; Gao, J.; Suga, H.; and Gerlt, J. *Proc Nat Acad Sci USA* **2014**, 42, 15066-15077. DOI: 10.1073/pnas.1411772111.

We conclude that (i) the hydrogen bond between the backbone amide of Ser-127 and O4 of the pyrimidine moiety contributes a modest factor (~10^2) to the 10^{17} rate enhancement and (ii) the stabilization of the anionic intermediate is accomplished by electrostatic interactions, including its proximity of Lys-72. These conclusions are in good agreement with predictions obtained from hybrid quantum mechanical/molecular mechanical calculations.

INTRODUCTION

Elucidation of the structural strategies by which enzymes achieve their large rate enhancements is essential for understanding the evolution of enzymatic catalysis as well as facilitating the design of enzymes to catalyze novel reactions. Typically, the possible strategies are identified by high-resolution X-ray structural analyses, in which a stable mimic of a reactive intermediate or rate-determining transition state is bound in the active site. Then, site-directed mutagenesis is used to generate substitutions, in which the important interactions are removed or altered, with kinetic analyses and structural studies of the mutant enzymes allowing the importance of the interaction to be assessed. Indeed, for nearly 30 y, this approach has been used to evaluate the importance of interactions involving amino acid side chains. However, structural studies of many enzymes reveal the presence of hydrogen-bonding interactions between a backbone amide group donor and a heteroatom acceptor in the substrate. Such an interaction can contribute to catalysis by increasing in strength as the basicity/proton affinity of the acceptor is increased as the reaction coordinate is traversed (1). Perhaps

the best known example of such an interaction is the oxyanion hole in serine proteases, in which the substrate peptide carbonyl group is converted to an anionic tetrahedral intermediate (2). However, the importance of the presumed enhanced hydrogen-bonding interaction between the oxyanion hole and the substrate/intermediate has not been directly evaluated because of the difficulty in mutating the backbone amide group donor and/or the substrate acceptor. OMP decarboxylase (OMPDC) provides another important example of a hydrogen bonding interaction between a backbone amide group and a hydrogen bond acceptor in the substrate that has been proposed to be important for stabilizing a reactive intermediate. OMPDC is a paradigm for understanding the structural features responsible for catalytic efficiencies of enzymes, because the reaction is cofactor-independent; also, the rate enhancement (~10^{17}) and proficiency (affinity for the transition state; ~10^{23} M) are among the largest for any enzyme-catalyzed reaction (3). The rate enhancement is a composite of substrate destabilization and intermediate stabilization (**Figure 10A**) (4–11). The pKa of the UMP product is reduced from 30 to 34 in solution to ≤22 in the active site, requiring significant stabilization of the vinyl anion intermediate (~14 kcal/mol) (12–14).

Others and we have characterized structure–function relationships in the OMPDC from *Methanobacter thermoautotrophicum* (MtOMPDC). The structure of MtOMPDC with the potent competitive inhibitor 6-hydroxyuridine 5′-monophosphate [BMP; also known as 1-(5′-phospho-β-D-ribofuranosyl)barbiturate] reveals (i) a hydrogen bonding interaction between the backbone amide group of Ser-127 and O4 of the pyrimidine

moiety and (ii) the proximity of carbon-6 to the ε-ammonium group of Lys-72, the general acidic catalyst. Accordingly, two structural strategies have been proposed for stabilization of the anionic intermediate (**Figure 10B**): (i) an enhanced hydrogen-bonding interaction between the Ser-127 backbone amide and O4 of the pyrimidine moiety of the intermediate through a carbene resonance structure (anionic charge on O4) (8, 11, 15) and (ii) electrostatic interactions between the anionic charge on carbon-6 with the e-ammonium group of Lys-72 (4, 9). Both are challenging to test: (i) substitutions for backbone amide groups are difficult to construct, and (ii) Lys-72 is the essential general acid catalyst. We constructed the S127P mutant to evaluate the importance of the interaction of the backbone amide group with O4; the mutant has a dramatically reduced value for k_{cat}/K_M (10^6-fold) compared with WT MtOMPDC. However, a comparison of liganded X-ray structures of the WT and S127P mutant revealed a displacement of BMP in the active site as the result of the steric interactions between the proline side chain and O4, thereby making interpretation of the importance of the interaction of the backbone amide group with O4 difficult (8). We now report kinetic constants for the variant of MtOMPDC in which the backbone amide group of Ser-127 is replaced with an ester group. The Ser to L-glycerate (^{HO}S) substitution (S127^{HO}S) eliminates the backbone hydrogen bond while keeping the interaction of the side chain hydroxyl group with N3 of the orotate moiety intact. However, the substitution juxtaposes the backbone ester oxygen with O4, having the potential to reduce the affinity for the substrate and/or destabilize the carbene-like resonance structure of the anionic intermediate. Using 5-fluoroOMP (FOMP) as the substrate, we

found that the value of k_{cat} is reduced ≤50-fold and that the value of K_M is increased ≤1.5-fold relative to WT MtOMPDC. We conclude that the hydrogen bond between the backbone amide and O4 contributes a factor of $\leq 10^2$ to the 10^{17} rate enhancement, suggesting that the anionic intermediate is stabilized primarily by electrostatic interactions within the active site, including Lys-72. These conclusions are in good agreement with predictions obtained from hybrid quantum mechanical/molecular mechanical calculations. They also illustrate the considerable power of combining a variety of biochemical and biophysical approaches (experimental enzymology, in vitro translation, crystallography, and computation) to tackle a challenging biochemical problem and the importance of the interactions of backbone amide groups of enzymes with substrates to stabilize reactive intermediates and transition states.

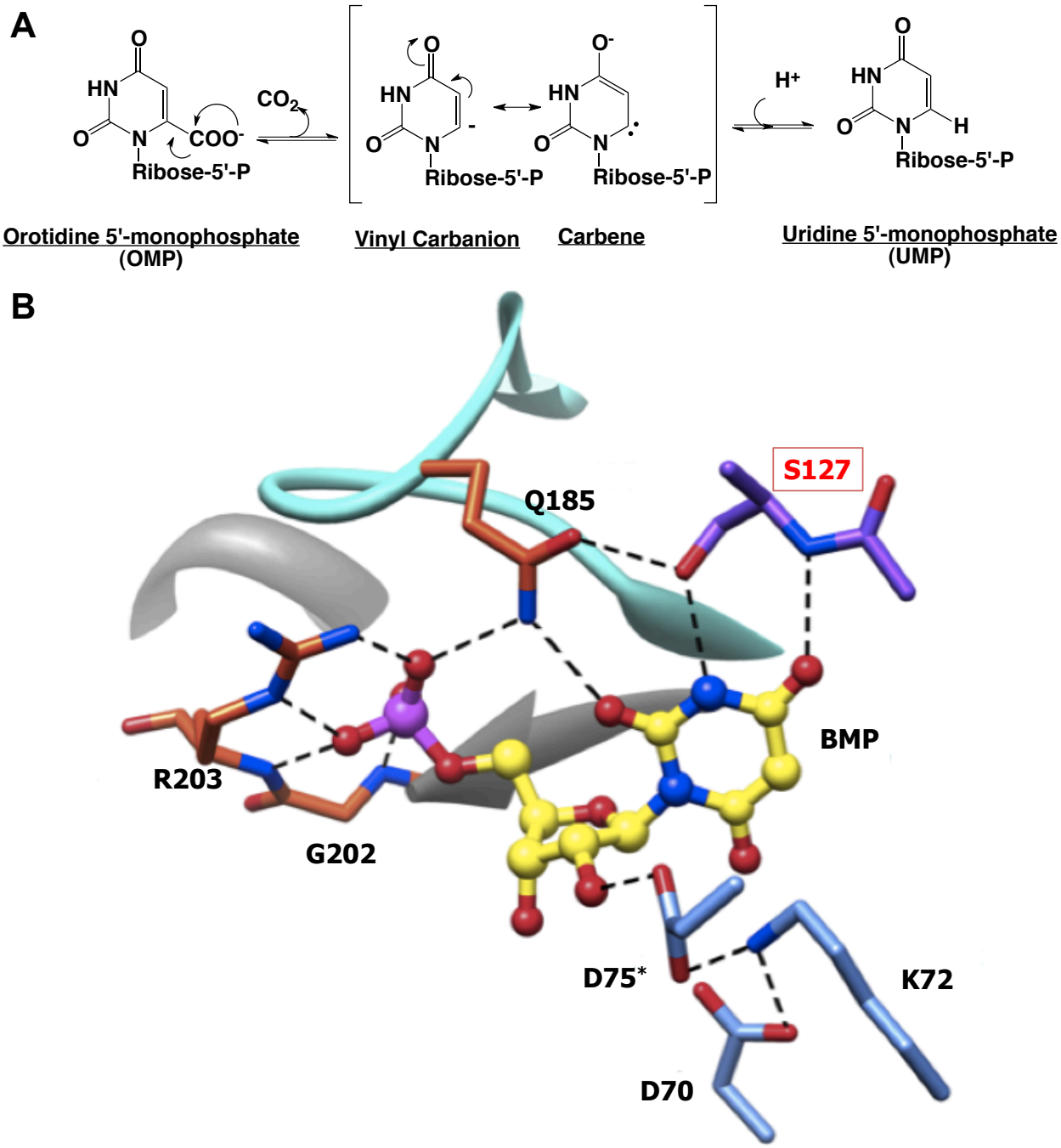


Figure 10. *(A)* Reaction catalyzed by OMPDC. *(B)* Active site of MtOMPDC with 6-OH uridine 5'-monophosphate (BMP), a vinyl anion intermediate analog. The figure was generated using coordinates from PDB ID 3LTP using Chimera (42).

RESULTS

In Vitro Protein Synthesis. We used the H128N mutant of MtOMPDC to eliminate the possibility of general base-catalyzed hydrolysis of the backbone ester bond by the spatially proximal imidazolium group of His-128. The kinetic constants for H128N are identical to those for WT enzyme; no structural perturbation is observed in the X-ray structure (**Appendix A, Figure 27 and Table 6**). In our experiments, we used the PURExpress In Vitro Translation Kit prepared from purified components from *Escherichia coli* (New England Biolabs). A yield of 0.22 mg/mL H128N was achieved by (i) using a synthetic gene optimized for codon use for E. coli and a low probability of forming secondary structures in the mRNA transcript (Genscript), (ii) optimizing the spacing between the ribosome binding site and the initiation codon, (iii) deleting the N-terminal 11 amino acids that are disordered in X-ray structure, and (iv) optimizing the amount of plasmid. The flexible tRNA acylation ribozyme system was used to acylate an orthogonal amber suppressor tRNA with the 3,5-dinitrobenzyl ester of ^{HO}S or Ser; acylation yields were ~40% (16–18) (**Appendix A**). Using suppressor tRNA charged with either ^{HO}S or Ser, full-length proteins with easily measurable activities were produced by in vitro translation of the template with the amber codon (TAG) replacing the WT codon for Ser (AGC; see below) (**Figure 11A**). In the absence of suppressor tRNA, a full-length protein with negligible activity (**Appendix A, Figure 28** reaction b, and **Table 7**) was produced as the result of adventitious suppression of the amber codon by glutamine (encoded by the CAG/CAA codon; S127Q; see below).

In all reactions, small quantities of the N-terminal residue 12– 126 peptide were formed (**Figure 11A**); this early termination may be explained by the incomplete acylation of the suppressor tRNA.

To verify that S127^{HO}S is produced with suppressor tRNA charged with ^{HO}S, reactions were incubated at pH 10 and 70 °C for 1 h to cleave the ester backbone (**Figure 11B**). The amount of full-length protein obtained with suppressor tRNA charged with Ser in the absence of suppressor tRNA was modestly decreased. In contrast, the

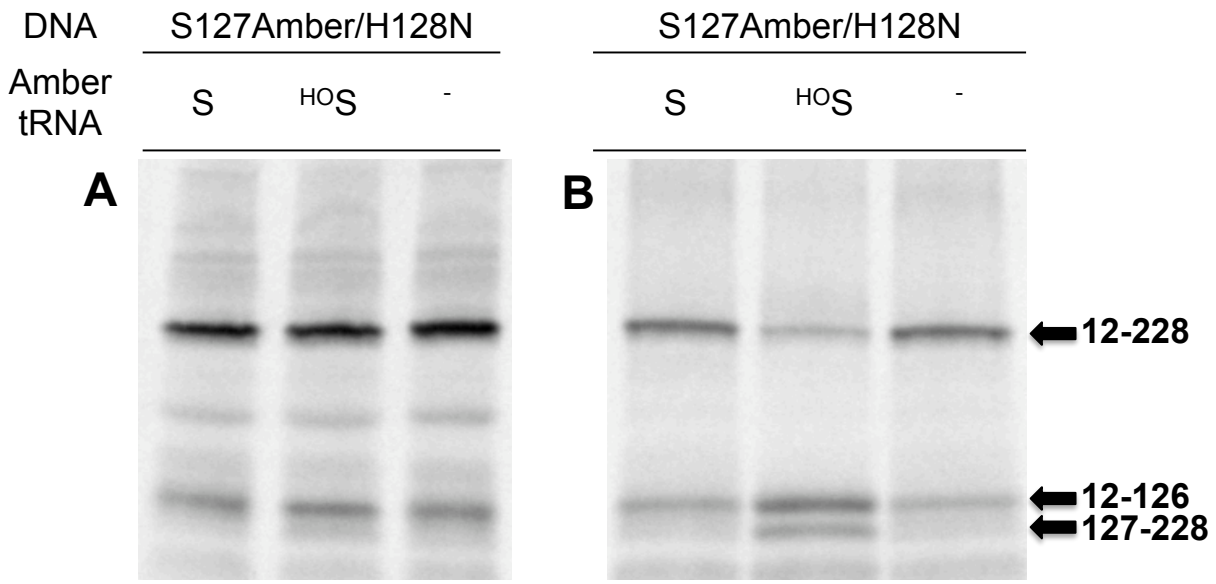


Figure 11. Autoradiograph of an SDS-PAGE gel of the *in vitro* translation reactions containing the S127amber/H128N plasmid with amber suppressor tRNA charged with L-serine (S); L-glycerate (^{HO}S); no amber tRNA (-). Panel A, no alkali or heat treatment prior to SDS-PAGE experiment. Panel B, reactions incubated at pH 10 and 70 °C for 1-hour prior to SDS-PAGE experiment. The expected protein fragments are marked on the right.

amount of full-length protein obtained with suppressor tRNA charged with ^{HO}S was significantly decreased; the bands corresponding to the N-terminal residue 12–126 and the C-terminal residue 127–228 peptides of the full-length protein were produced in the expected 2:1 ratio (4 and 2 Met residues, respectively), confirming the presence of the backbone ester. The reaction components precipitate under these conditions, and therefore, enzymatic activities could not be used to verify the expected loss of activity for the S127^{HO}S substitution (activity was lost for the protein produced when the suppressor tRNA was charged with either ^{HO}S or Ser).

MS Analysis of in Vitro-Synthesized Proteins. MS/MS analysis of SDS/PAGE-purified and Glu-C endopeptidase-digested in vitro expressed WT (H128N) identified a peptide with a mass of 1,422.64 Da in high abundance, corresponding to VFLLTEM$_{ox}$SNPGAE (M$_{ox}$ is oxidized Met; residues 120–132) (**Figure 12A**). A similar analysis of S127^{HO}S identified a peptide with a mass of 1,423.66 Da in high abundance, corresponding to VFLLTEM$_{ox}$(S→^{HO}S)NPGAE (residues 120–132) (**Figure 12B**). The 1-Da difference in the masses of the peptide fragments from WT and S127^{HO}S is equal to the difference in the masses of serine and glycerate. In addition, the peptide fragment corresponding to M$_{ox}$SNPGAEMoxFIQGAADE (residues 126–141) was detected in high abundance in the WT protein (**Appendix A, Figure 29A**). An analogous peptide was detected in S127^{HO}S, (S→^{HO}S)NPGAEM$_{ox}$FIQGAADE, except that it lacks Met-126 (**Appendix A, Figure 29B**). This fragment is not the product of Glu-C hydrolysis (Glu-C cleaves after glutamate or aspartate residues) and presumably, is formed by

hydrolysis of the labile ester bond between Met-126 and ^{HO}S 127 during gel staining/destaining process and before performing the MS analysis. Thus, the detection of this fragment provides strong evidence for the incorporation of ^{HO}S at residue 127.

The amino acid incorporated in the absence of suppressor tRNA was identified as glutamine (i.e., S127Q) (**Appendix A, Figure 30**). Analysis of the protein produced in the Ser-suppressed control showed the presence of peptides containing both Ser and Gln at the 127th position (WT and S127Q) (**Appendix A, Figure 31**). Analysis of the ^{HO}S-suppressed control detected the presence of only ^{HO}S at the 127th position, suggesting the production of a negligible amount of S127Q when the ^{HO}S-charged suppressor tRNA was used. This conclusion is also supported by the nearly complete hydrolysis of the full-length S127^{HO}S protein (**Figure 11**), suggesting that the majority of the protein produced in the presence of suppressor tRNA charged with L-glycerate contains an alkali–labile ester linkage.

Kinetic Analysis of in vitro-Synthesized Proteins. The in vitro translation reactions were assayed using FOMP, because it is ~500-fold more reactive than OMP due to the electron-withdrawing substituent that stabilizes the anionic intermediate (19). In this way, the values of k_{cat}, K_M, and k_{cat}/K_M could be determined by following the reactions to completion for the in vitro-expressed WT (H128N), WT (H128N) expressed in the Ser-suppressed reaction, and S127^{HO}S. The kinetic constants are displayed in **Table 1**. WT (H128N) produced by in vitro translation using a plasmid template with the Ser codon has values of k_{cat} and K_M similar to those for the WT

(H128N) expressed in and purified from E. coli, confirming the integrity of in vitro-produced enzyme. The protein produced in the absence of a charged suppressor tRNA (S127Q; see above), the expected contaminant in reactions using suppressor tRNA charged with Ser or ^{HO}S, represents the upper limit of contaminating activity present in proteins produced by amber suppression using amber tRNA; this activity is negligible (**Appendix A**, **Figure 28** reaction b and **Table 7**). This control rules out the possibility that the activity measured for either the WT protein produced by suppression with Ser or S127^{HO}S produced by suppression with ^{HO}S is the result of a catalytically active contaminant produced by the components of the in vitro translation reaction.

Furthermore, as described above, we observed that Gln is incorporated in the absence of suppressor tRNA. This incorporation is easily explained: the amber codon is TAG, and the codons for Gln are CAG/CAA. We prepared the S127Q mutant from a strain of E. coli in which the gene encoding OMPDC is disrupted: the value of k_{cat}/K_M for the S127Q is 2.4 X 10^2 M^{-1} s^{-1}, significantly lower than those measured for the proteins produced by suppressor tRNA charged with either ^{HO}S (S127^{HO}S) or Ser (**Appendix A**, **Table 8**). The value of k_{cat} for WT (H128N) produced using suppressor tRNA charged with Ser is twofold less than that measured for WT (H128N) produced using a plasmid template with the Ser codon; this decrease can be explained by the presence of the adventitious, catalytically inactive S127Qdetected in our MS analysis (**Appendix A, Figure 31**). The value of k_{cat} for S127^{HO}S produced using suppressor tRNA charged with ^{HO}S (6.7 s^{-1}) is reduced 50-fold from WT (H128N) produced using a plasmid

template with the Ser codon and 25-fold from the WT (H128N) produced using suppressor tRNA charged with Ser (**Table 1**). Our MS analysis of $S127^{HO}S$ did not detect S127Q, and therefore, the upper limit on the reduction in activity associated with the amide to ester substitution is 50-fold; assuming that $S127^{HO}S$ contains the same amount of S127Q as the WT control, the reduction in activity is 25-fold. Irrespective, we conclude that the activity of $S127^{HO}S$ is significant and associated with the amide to ester substitution. The value of KM for $S127^{HO}S$ (140 μM) is elevated threefold from that measured for WT (H128N) produced using a plasmid template with the Ser codon and identical, within error, to that measured for WT produced using suppressor tRNA charged with Ser. We conclude that binding of the substrate is not significantly perturbed by backbone ester–O4 interaction that results from the amide to ester substitution.

Table 1. Kinetic constants for FOMP decarboxylation, measured at 25°C and pH 7.1.

	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} M^{-1})
H128N[a]	340 ± 40	90 ± 10	(3.8 ± 0.6) x 10^6
H128N[b]	280 ± 20	52 ± 4	(5.4 ± 0.6) x 10^6
H128N[c]	140 ± 40	120 ± 20	(1.2 ± 0.4) x 10^6
$S127^{HO}S$[c]	6.7 ± 1.9	140 ± 20	(4.8 ± 1.5) x 10^4

[a], purified from *E. coli*; [b], *in vitro* translation using wild type plasmid; [c], *in vitro* translation using amber suppression.

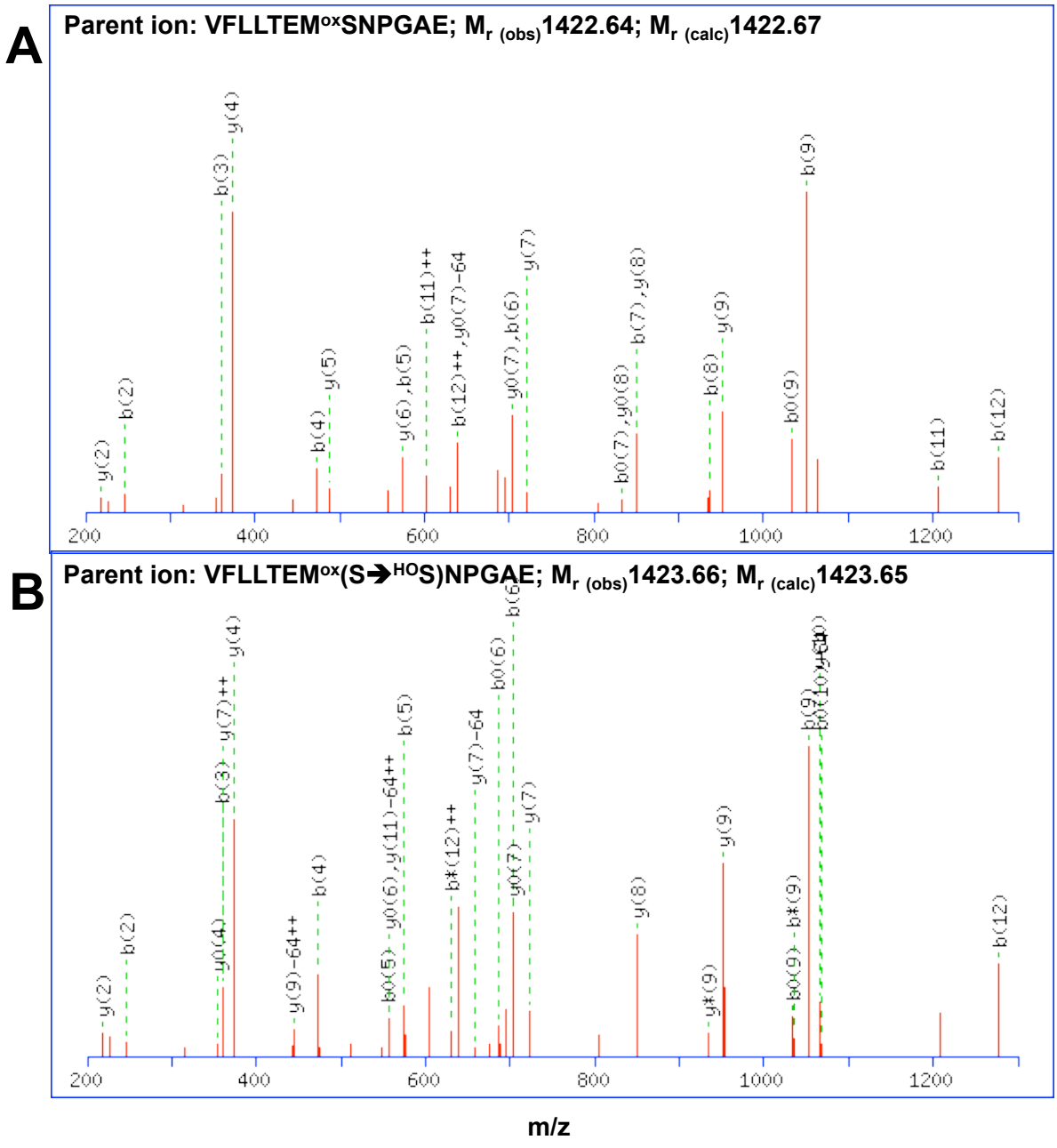


Figure 12. MS/MS analysis of in vitro-synthesized WT (H128N) and S127^{HO}S proteins. (A) MS/MS fragmentation of the parent ion from the WT protein corresponding to the sequence VFLLTEM$_{ox}$SNPGAE (residues 120–132). (B) MS/MS fragmentation of the parent ion from the S127^{HO}S protein corresponding to the sequence VFLLTEM$_{ox}$(S→^{HO}S)NPGAE (residues 120–132). The observed [$M_{r\ (obs)}$] and calculated [$M_{r\ (calc)}$] monoisotopic masses of the parent ions are indicated on each spectrum. Assigned b- and y-fragment ions are numbered according to the modified Roepstorff and Fohlman nomenclature (31, 32).

Hybrid Quantum Mechanical and Molecular Mechanical Studies. We carried out molecular dynamics simulations using a combined quantum mechanical and molecular mechanical (QM/MM) potential to dissect the effects of the backbone ester substitution on catalysis. The approach has been validated previously and applied to the WT MtOMPDC (5). The free energy profiles for WT and S127^{HO}S mutant are depicted in **Figure 13A**. The free energy barrier was calculated to be 4.3 kcal/mol greater for S127^{HO}S than for the WT enzyme. The charge distribution of the carbanion intermediate of the decarboxylation reaction is highly delocalized over the entire pyrimidine ring in the active site of OMPDC, with an average partial atomic charge of only −0.22 a.u. on the C6 position from Mulliken population analysis as well as electrostatic potential fitting (**Appendix A**, **Tables 9 and 10**). The remaining charge density is distributed to the C5H and C4O4 groups, with net charges of −0.36 and −0.26 a.u., respectively, and the rest is distributed to the nitrogen sites. Although this seems to be consistent with a carbene-like structure at the C6 position, its relatively small net charge is a result of combined effects of π-conjugation and σ-delocalization (evidenced by the negative charge at C5). In contrast, the anionic charge is primarily localized on the carboxylate group of OMP in the Michaelis complex state. Thus, the relief of electrostatic stress of the charge-localized reactant state, caused by interaction with the Asp residue in the active site, provides a significant contribution to the rate acceleration. The S127^{HO}S substitution has the largest effects on the carbanion intermediate, with a net electron density shift of 0.03 a.u. from the O4 carbonyl oxygen to the C6 atom through π-conjugation compared with the WT OMPDC. We attribute this charge shift to the

negative impact of the lone pair electron density of the ester oxygen in the S127^{HO}S mutant. Nevertheless, the effect of the backbone mutation on the charge redistribution of the pyrimidine ring is relatively small, consistent with the observation that rate reduction is modest. The S127^{HO}S substitution induces significant distortion of the active site at the transition state (**Figure 13B**). In WT, the O4 oxygen of the orotate moiety of the substrate is hydrogen-bonded to the Ser-127 backbone amide, with an average donor–acceptor distance of 3.0 Å. However, the amide to ester substitution introduces an electrostatic clash with the O4, elongating the distance between O4 and the ester oxygen by 0.5 Å to an average of 3.5 Å (**Figure 13B** and **Appendix A**, **Figure 32**).

We performed interaction energy decomposition analysis between the orotate moiety and residue 127 to estimate the energetic consequence as a result of the Ser to ^{HO}S backbone substitution (**Appendix A**, **Table 11**). In WT, the Ser-127 backbone amide contributes 1.9 kcal/mol to transition state stabilization relative to that of the Michaelis complex. However, in S127^{HO}S, the interaction between orotate and the ^{HO}S backbone ester group results in a destabilizing effect of 1.4 kcal/mol in going from the reactant state to the transition state as the anionic charge is shifted from the carboxylate group into the pyrimidine ring. Therefore, in the WT enzyme, this charge redistribution is stabilized by the Ser-127 backbone amide hydrogen bond, whereas in S127^{HO}S, its formation is penalized by interactions with the ester oxygen atom lone pairs.

Discussion

We have successfully used in vitro translation to produce sufficient quantities of the $S127^{HO}S$ substitution of MtOMPDC to allow experimental assessment of the importance of the hydrogen bond between the backbone amide group of Ser-127 and O4 of the pyrimidine moiety of the substrate. Interpretation of the experimental results together with parallel computational predictions of the effects of the substitution on the reaction coordinate allow important conclusions about the structural strategies for stabilization of the anionic intermediate. The values of both k_{cat} and k_{cat}/K_M for WT MtOMPDC are partially diffusion-controlled (i.e., the transition states for substrate binding and product dissociation are similar in energy to that for decarboxylation). As a result, the value of k_{cat} underestimates the rate constant for decarboxylation of the WT•FOMP complex. Based on both the fractional inverse dependence of k_{cat} on solvent viscosity (slope = 0.6) and a comparison of the values of k_{cat} for both FOMP and OMP for the D70N mutant, where chemistry is rate-determining, we estimate that the rate constant for decarboxylation for the WT•FOMP complex is ~1,000 s^{-1} (19). Therefore, the ester substitution reduces the rate of decarboxylation of the $S127^{HO}S$•FOMP complex ~150-fold. This rate reduction corresponds to an increase of ~3 kcal/mol in the activation energy barrier, which is in good agreement with the computed increase of 4.3 kcal/mol in the free energy barrier for $S127^{HO}S$ relative to the WT enzyme. In contrast, the value of K_M is increased by only twofold, suggesting that the juxtaposition of O4 with the ester bond does not significantly interfere with the binding of FOMP.

Our kinetic data and calculations together establish that stabilization of the vinyl anion intermediate is decreased ~3 kcal/mol by a combined effect of loosening the hydrogen-bonding interaction of O4 with the Ser-127 backbone amide and juxtaposing the O4 and ester functional groups. Interaction energy decomposition analysis shows that, of the ~3-kcal/mol overall increase in the activation barrier because of the $S127^{HO}S$ substitution, <2 kcal/mol may be attributed to the stabilizing effects of the hydrogen bond between O4 and the Ser-127 amide group. This energy change represents only a small contribution to the overall catalytic effect that lowers the free energy barrier of the uncatalyzed process in water by 23 kcal/mol. Because development of the O4 anion resonance structure is essential for stabilization of a carbene-like intermediate, the relatively small stabilization effects by the Ser-127 backbone amide suggest that it is unlikely to be a major resonance form for the decarboxylation intermediate. Therefore, in the context of the 1017 rate enhancement (23-kcal/mol decrease in $\Delta G^{\ddagger}$), the interaction between the backbone amide and O4 provides one of several modest contributions; others include substrate destabilization by enforced proximity of the substrate carboxylate group to Asp-70 in an otherwise hydrophobic pocket (5, 8).

We conclude that the dramatically reduced value for k_{cat}/K_M (10^6-fold) observed for the S127P mutant (8) is dominated by altered binding of the substrate/transition state/intermediate to the active site, resulting in suboptimal geometry for decarboxylation of the OMP substrate and/or protonation of the anion intermediate to

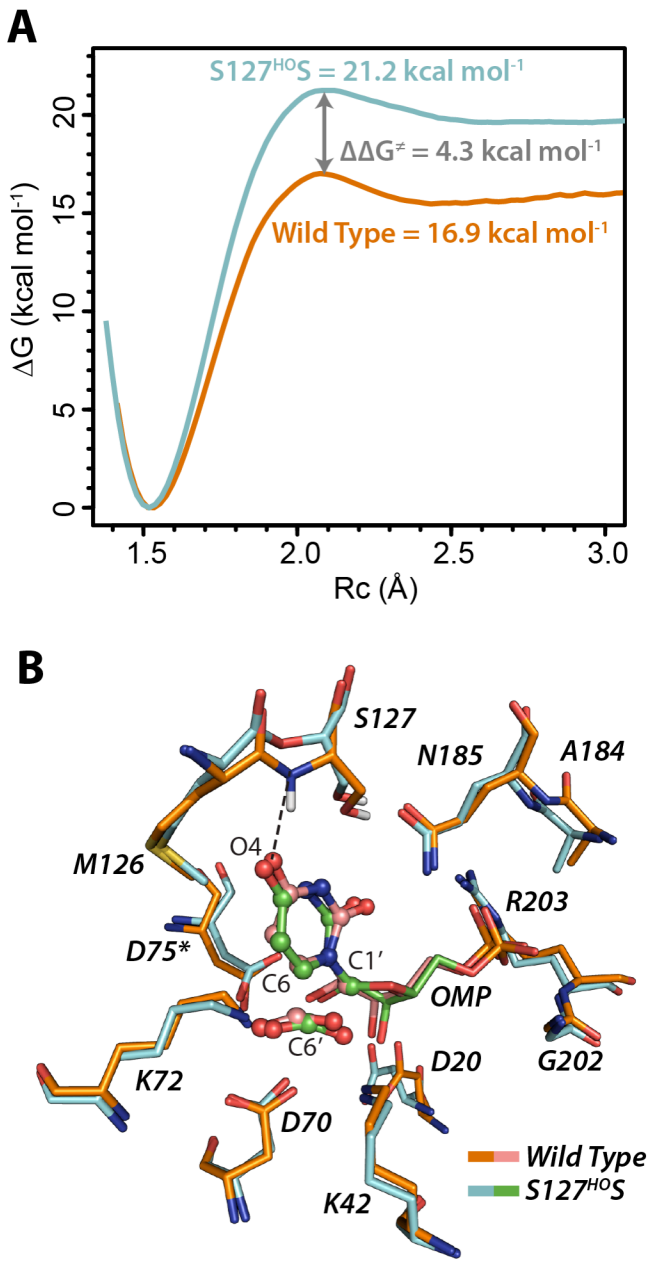


Figure 13. (A) Computed free energy profile for the decarboxylation reaction in wild type MtOMPDC and S127^{HO}S. The reaction coordinate is defined by the C6-C6′ distance. (B) representative structures of the TS ensemble from the MD simulations for the wild-type enzyme (backbone colored) and the S127^{HO}S mutant (backbone in brown). The QM region is shown as ball-and-stick.

yield the UMP product. Also, although electrostatic interactions between C6 of the intermediate and the ε-ammonium group of Lys-72, the general acid catalyst, have been proposed to be the major contributor to the ~14-kcal/mol stabilization of the intermediate, the delocalized charge calculated for the intermediate suggests that this interaction is not the major contributor; instead, relief of the destabilizing interactions between the localized charge on the orotate carboxylate group with Asp-70 relative to the charge delocalized intermediate must be important for achieving the rate enhancement. In contrast to prior studies of MtOMPDC reported by these laboratories, in which X-ray structures were determined for all mutant proteins (a total of 75 Protein Data Bank depositions), we were unable to obtain sufficient amounts of $S127^{HO}S$ for structural characterization because of the costs associated with scaling up the in vitro translation reactions. However, the robust activity using FOMP provides compelling support for the validity of the experimental results and their interpretations. Finally, the success of these experiments shows that in vitro translation can be used to generate sufficient quantities of backbone amide to ester substitutions to test the importance of their interactions with substrates/intermediates/transition states interactions in other enzymes (e.g., serine proteases, haloalkane dehalogenases, and enoyl-CoA hydratase) (2, 20, 21).

Materials and Methods

OMP and FOMP were prepared using previously published chemoenzymatic synthesis (4, 19, 22). Cell free protein synthesis kits, murine RNase inhibitor, T7 RNA polymerase, NTPs, Taq polymerase, and dNTPs were purchased from New England Biolabs.

In Vitro Protein Synthesis. In vitro protein synthesis was performed using the PURExpress In Vitro Protein Synthesis Kit from New England Biolabs (23). Release factor-1 was omitted in all of the reactions unless stated otherwise (24, 25). Murine RNase inhibitor was added to all reactions at a concentration of 0.8 U/μL. The concentration of template plasmid used was 5 ng/μL. The MtOMPDC H128N plasmid was used for the synthesis of WT protein. MtOMPDC S127Amber/H128N plasmid was used for in vitro protein synthesis of the S127^{HO}S. The flexible tRNA acylation ribozyme system was used to acylate an orthogonal amber suppressor tRNA with the 3,5-dinitrobenzyl ester of ^{HO}S or Ser. Acylation yields were ~40% (16, 17, 26) (**Appendix A**); 100 pmol amber suppressor tRNA charged with Ser or ^{HO}S was added per 1 μL in vitro protein synthesis reaction. After mixing the components, the reactions were incubated at 37 °C for 4 h. After 4 h, the reactions were brought to 4 °C by placing the tubes on ice. Radioisotope labeling of in vitro-synthesized proteins was achieved by adding [^{35}S]-methionine at a final concentration of 10 nCi/μL.

FOMP Decarboxylation Assay. The FOMP decarboxylation assays were performed

using a continuous UV spectrophotometric assay as previously described (18, 19, 27) with modifications. All assays were performed at 25 °C in 10 mM MOPS [3-(N-morpholino)-propane sulfonic acid] buffer and 100 mM NaCl (pH 7.1) in a 1,000-µL quartz cuvette with a path length of 10 mm; a Cary Bio 300 spectrophotometer was used for all measurements. The concentration of the FOMP stock solution was determined by measuring absorbance at 270 nm of a known dilution in 0.1 M HCl using the molar extinction coefficient of FOMP (ε_{FOMP} = 9,895 M^{-1} cm^{-1}). Concentrations of proteins purified from E. coli (MtOMPDC H128N and S127Q) were determined by measuring absorbance at 280 nm and 25 °C in 10 mM MOPS and 100 mM NaCl (pH 7.1) using the molar extinction coefficient (ε_{280} = 5,960). Final enzyme concentrations of 30 nM and 3.5 µM were used for H128N and S127Q, respectively. The concentration of in vitro synthesized full-length protein was determined as described elsewhere (**Appendix A**). In vitro synthesized proteins were assayed by adding 5- or 10-µL aliquots of the in vitro reaction mixture in the assay, as is indicated in **Appendix A**, **Table 7**. The decay in absorbance corresponding to the disappearance of FOMP/formation of 5-fluoro uridine monophosphate was monitored at 300 nm **(Appendix A**, **Figure 28**). The value of k_{cat} was determined using saturating concentrations of FOMP and measuring the initial rate of decarboxylation (V_o; corresponding to ≤5% reaction) using the rate of change in absorbance and $\Delta\varepsilon_{300}$ of −489 M^{-1} cm^{-1}. This rate represents the V_{max}. The k_{cat} was calculated by dividing V_{max} with the enzyme concentration determined previously; the value of K_M was determined by following the first-order decay of decarboxylation at low FOMP concentrations (≤K_M).

The decay curve was fit to a first-order decay function using nonlinear regression. The rate constant generated from this model represents Vmax/KM, from which the K_M was calculated using the enzyme concentration and k_{cat} determined previously.

MS Analysis of in Vitro-Synthesized Proteins. Protein bands corresponding to full-length OMPDC (**Appendix A**, **Figure 33**) were excised using a scalpel, placed in a clean Eppendorf tube, and stored at −20 °C. MS analysis was conducted using a Thermo LTQ Velos ETD Pro Mass Spectrometer. Before liquid chromatography MS/MS analysis, each protein band in its gel slice was crushed, destained, and dehydrated in 50% (vol/vol) acetonitrile containing 25 mM ammonium acetate. Glu-C digestion was performed using Staph Protease–Sequencing Grade at 1:10 (Worthington Biochemical) using a CEM Discover Microwave Reactor for 15 min at 55 °C at 50 W. The digested peptides were extracted three times using 50% (vol/vol) acetonitrile containing 5% (vol/vol) formic acid, pooled, and dried using a Speedvac (Thermo Scientific). The dried peptides were suspended in 5% (vol/vol) acetonitrile containing 0.1% formic acid and applied to the liquid chromatography. Control and data acquisition of the mass spectrometer were done using Xcalibur 2.2 under the data-dependent acquisition mode; after an initial full scan, the top five most intense ions were subjected to MS/MS fragmentation by collision-induced dissociation. The raw data were processed by Mascot Distiller (Matrix Sciences) and then Mascot (version 2.4). The result was searched against the MtOMPDC H128N sequence and the National Center for Biotechnology Information NR Protein database.

Hybrid QM/MM Studies. Combined QM/MM molecular dynamics simulations were performed on an OMP–MtOMPDC complex constructed on the basis of the crystal structure of the BMP-bound complex (Protein Data Bank ID code 1LOR) following the procedure reported previously (5). The SQUANTUM implementation of the AM1 method (28) in CHARMM (29) (version c35a1) was used to model the QM region, which encompassed the entire orotate moiety. Additional details are provided in **Appendix A**.

ACKNOWLEDGEMENTS

We thank Prof. Kendall N. Houk for valuable discussions; Dr. Corinna Tuckey (New England BioLabs) for helpful advice and materials for in vitro translation; Dr. Susan Martinis and Mr. Aaron Frimel (University of Illinois at Urbana–Champaign) for assistance in performing the radioactive isotope experiments; and Dr. Peter Yau and Dr. Brian Imai (Roy J. Carver Biotechnology Center, University of Illinois at Urbana–Champaign) for help with the MS analysis. We also thank the Minnesota Supercomputing Institute for computational resources. This research was supported by National Institutes of Health Grants R01GM065155 (to J.A.G.) and R01GM046367 (to J.G.).

REFERENCES

(1) Cleland WW, Frey PA, Gerlt JA (1998) The low barrier hydrogen bond in enzymatic catalysis. J Biol Chem 273(40):25529–25532.

(2) Bryan P, Pantoliano MW, Quill SG, Hsiao HY, Poulos T (1986) Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. Proc Natl Acad Sci USA 83(11):3743–3745.

(3) Radzicka A, Wolfenden R (1995) A proficient enzyme. Science 267(5194):90–93.

(4) Van Vleet JL, Reinhardt LA, Miller BG, Sievers A, Cleland WW (2008) Carbon isotope effect study on orotidine 5′-monophosphate decarboxylase: Support for an anionic intermediate. Biochemistry 47(2):798–803.

(5) Wu N, Mo Y, Gao J, Pai EF (2000) Electrostatic stress in catalysis: Structure and mechanism of the enzyme orotidine monophosphate decarboxylase. Proc Natl Acad Sci USA 97(5): 2017–2022.

(6) Callahan BP, Wolfenden R (2004) OMP decarboxylase: An experimental test of electrostatic destabilization of the enzyme-substrate complex. J Am Chem Soc 126(45): 14698–14699.

(7) Miller BG, Snider MJ, Wolfenden R, Short SA (2001) Dissecting a charged network at the active site of orotidine-5′-phosphate decarboxylase. J Biol Chem 276(18):15174–15176.

(8) Iiams V, et al. (2011) Mechanism of the orotidine 5′-monophosphate decarboxylase catalyzed reaction: Importance of residues in the orotate binding site. Biochemistry 50(39):8497–8507.

(9) Chan KK, et al. (2009) Mechanism of the orotidine 5′-monophosphate decarboxylase catalyzed reaction: Evidence for substrate destabilization. Biochemistry 48(24):5518-5531.

(10) Fujihashi M, et al. (2013) Substrate distortion contributes to the catalysis of orotidine 5′-monophosphate decarboxylase. J Am Chem Soc 135(46):17432–17443.

(11) Houk KN, Lee JK, Tantillo DJ, Bahmanyar S, Hietbrink BN (2001) Crystal structures of orotidine monophosphate decarboxylase: Does the structure reveal the mechanism of nature's most proficient enzyme? ChemBioChem 2(2):113–118.

(12) Toth K, et al. (2007) Product deuterium isotope effect for orotidine 5′-monophosphate decarboxylase: Evidence for the existence of a short-lived carbanion intermediate. J Am Chem Soc 129(43):12946–12947.

(13) Amyes TL, Wood BM, Chan K, Gerlt JA, Richard JP (2008) Formation and stability of a vinyl carbanion at the active site of orotidine 5′-monophosphate decarboxylase: pKa of the C-6 proton of enzyme-bound UMP. J Am Chem Soc 130(5):1574–1575.

(14) Toth K, et al. (2010) Product deuterium isotope effects for orotidine 5′-monophosphate decarboxylase: Effect of changing substrate and enzyme structure on the partitioning of the vinyl carbanion reaction intermediate. J AmChem Soc 132(20):7018–7024.

(15) Lee JK, Houk KN (1997) A proficient enzyme revisited: The predicted mechanism for orotidine monophosphate decarboxylase. Science 276(5314):942–945.

(16) Ohta A, Murakami H, Higashimura E, Suga H (2007) Synthesis of polyester by means of genetic code reprogramming. Chem Biol 14(12):1315–1322.

(17) Goto Y, Katoh T, Suga H (2011) Flexizymes for genetic code reprogramming. Nat Protoc 6(6):779–790.

(18) Amyes TL, Richard JP, Tait JJ (2005) Activation of orotidine 5′-monophosphate decarboxylase by phosphite dianion: The whole substrate is the sum of two parts. J Am Chem Soc 127(45):15708–15709.

(19) Wood BM, Chan KK, Amyes TL, Richard JP, Gerlt JA (2009) Mechanism of the orotidine 5′-monophosphate decarboxylase-catalyzed reaction: Effect of solvent viscosity on kinetic constants. Biochemistry 48(24):5510–5517.

(20) Xu D, et al. (2004) QM/MM studies of the enzyme-catalyzed dechlorination of 4-chlorobenzoyl-CoA provide insight into reaction energetics. J AmChem Soc 126(42):13649–13658.

(21) Meriläinen G, Poikela V, Kursula P, Wierenga RK (2009) The thiolase reaction mechanism: The importance of Asn316 and His348 for stabilizing the enolate intermediate of the Claisen condensation. Biochemistry 48(46):11011–11025.

(22) Wood BM, et al. (2010) Conformational changes in orotidine 5′-monophosphate decarboxylase:"Remote" residues that stabilize the active conformation. Biochemistry 49(17):3514–3516.

(23) Shimizu Y, et al. (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19(8):751–755.

(24) Johnson DB, et al. (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol 7(11):779–786.

(25) Johnson DB, et al. (2012) Release factor one is nonessential in Escherichia coli. ACS Chem Biol 7(8):1337–1344.

(26) Murakami H, Ohta A, Ashigai H, Suga H (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis. Nat Methods 3(5):357–359.

(27) Desai BJ, et al. (2012) Conformational changes in orotidine 5′-monophosphate decarboxylase: A structure-based explanation for how the 5′-phosphate group activates the enzyme. Biochemistry 51(43):8665–8678.

(28) Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) Development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model. J Am Chem Soc 107(13):3902–3909.

(29) Brooks BR, et al. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4(2):187–217.

(30) Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25(13):1605–1612.

(31) Roepstorff P, Fohlman J (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed Mass Spectrom 11(11):601.

(32) Biemann K (1990) Appendix 5. Nomenclature for peptide fragment ions (positive ions). Methods Enzymol 193:886–887.

CHAPTER 3: CONFORMATIONAL CHANGES IN OROTIDINE 5'-MONOPHOSPHATE DECARBOXYLASE: A STRUCTURE-BASED EXPLANATION FOR HOW THE 5'-PHOSPHATE GROUP ACTIVATES THE ENZYME

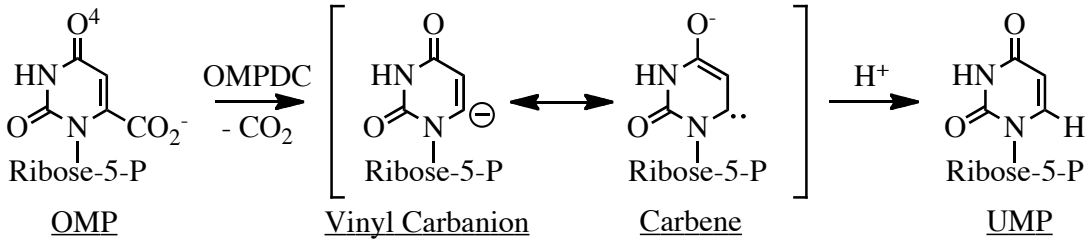
ABSTRACT: The reaction catalyzed by orotidine 5'-monophosphate decarboxylase (OMPDC) continues to interest the mechanistic enzymology community because the structural strategy used to achieve the large values of its rate enhancement (k_{cat}/k_{non} = 7.1 x 10^{16}) and catalytic proficiency [$(k_{cat}/K_M)/k_{non}$ = 4.8 x 10^{22} M] remain poorly understood. For all structurally characterized OMPDCs, the binding of a ligand (intermediate/transition state analog) is accompanied by a conformational change from an open, inactive conformation (E_o) that predominates in the absence of a ligand to a closed, active conformation (E_c) when the ligand is bound in the solvent-sequestered active site. As the substrate is bound (initially as the E_o•S complex that undergoes a conformational change to the E_c•S complex) and traverses the reaction coordinate to form the stabilized vinyl carbanion/carbene intermediate, interactions are enforced that destabilize the carboxylate group of the substrate as well as stabilize the intermediate (in the E_c•$S^{‡}$ complex).

Focusing on the OMPDC from *Methanothermobacter thermautotrophicus* (MtOMPDC), the "remote" 5'-phosphate group of the substrate activates the enzyme 2.4 x 10^8-fold [comparing the values of k_{cat}/K_M for OMP and 1-(β-D-erythrofuranosyl)orotic acid (EO; the 5'-hydroxymethylphosphate truncated analog of OMP)]; the activation is equivalently described by an intrinsic binding energy (IBE) of 11.4 kcal/mol for the 5'-phosphate group. We studied the roles of residues in the activation that 1) directly contact the 5'-phosphate group (Gln 185 and Arg 203); 2) participate in a hydrophobic cluster near the base of the closed active site loop that sequesters the bound substrate from solvent (Val 182; "remote" from the 5'-phosphate group); and 3) form hydrogen-bonding interactions across the interface between the "mobile" and "fixed" half-barrel domains of the $(\beta/\alpha)_8$-barrel structure (Thr 159, Arg 160, and Tyr 206; also "remote" from the 5'-phosphate group). Our structural and kinetic data support a structure-based model for the activation in which the IBE provided by the interaction of the 5'-phosphate group with its binding site is used, in part, to organize the active site loop and thereby enable interactions both near the N-terminus of the active site loop and across the domain interface that stabilize both the $E_c \bullet S$ and $E_c \bullet S^{\ddagger}$ complexes relative to the $E_o \bullet S$ complex, thereby a reduction of 23 kcal/mol in the value of $\Delta G^{\ddagger}$. The conclusion that the IBE of the "remote" 5'-phosphate group provides stabilization of both the $E_c \bullet S$ and $E_c \bullet S^{\ddagger}$ complexes, not just the $E_c \bullet S^{\ddagger}$ complex, is central to understanding the structural origins of enzymatic catalysis as well as the requirements for the de novo design of enzymes that catalyze novel reactions.

INTRODUCTION

Orotidine 5'-monophosphate decarboxylase (OMPDC;[1] **Scheme 1**) that catalyzes the final step in pyrimidine biosynthesis is a remarkable catalyst: the rate enhancement (k_{cat}/k_{non} = 7.1 x 10^{16}) and the catalytic proficiency [$(k_{cat}/K_M)/k_{non}$ = 4.8 x 10^{22} M]; affinity for the transition state) are among the largest measured for enzyme-catalyzed reactions (2, 3).

Scheme 1



We, and others (4-7), have been interested in defining the strategies used by OMPDC to reduce the value of $\Delta G^{\ddagger}$ by 23 kcal/mol. The reaction coordinate involves a vinyl carbanion/carbene intermediate that is stabilized by ≥14 kcal/mol by its interactions with the active site (8-10).

MtOMPDC, like all OMPDCs, is an obligate dimer, with the active sites located at the dimer interface. The interactions of conserved active site residues in the OMPDC from *Methanothermobacter thermautotrophicus*, MtOMPDC, the focus of this article, with BMP [l-(5-phospho-β-ribofuranosyl)barbituric acid], an intermediate/transition state analog (11), are shown in **Figure 14**. Each active site contains an essential Asp-Lys-Asp motif involving Asp 70 and Lys 72 from one polypeptide and Asp 75* from the symmetry related polypeptide as well as hydrogen-bonding interactions between the pyrimidine ring and both the OH group and the backbone amide group of Ser 127.

Two structural strategies can be envisioned for destabilization of the OMP substrate as the transition state is formed: 1) electrostatic/steric destabilization of the carboxylate group by the proximal Asp (Asp 70) (12, 13); and 2) desolvation of the carboxylate group by a proximal hydrophobic cavity (formed by Ile 96, Leu 123, Val 155, and Pro 180 (14, 15). The importance of these interactions has been tested; each contributes ~4 kcal/mol to the reduction in $\Delta G^{\ddagger}$.

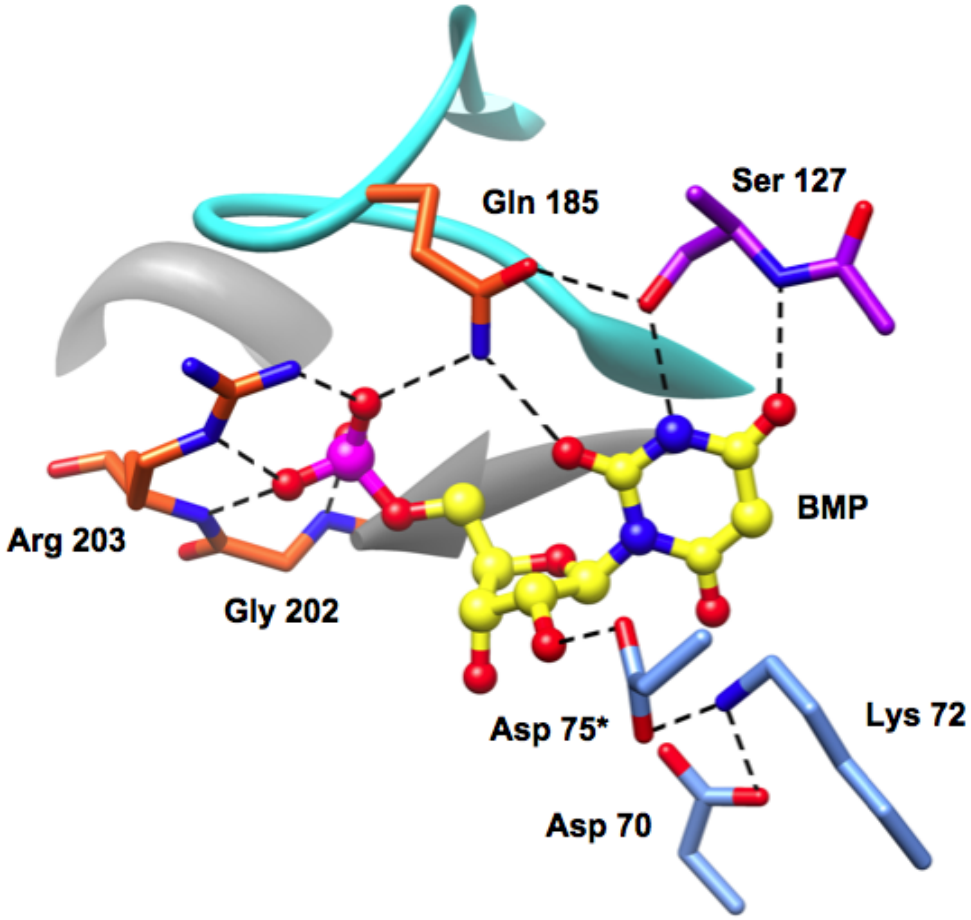


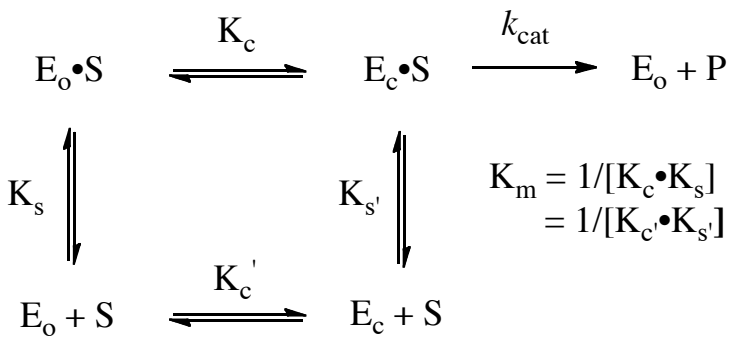
Figure 14. Active site of MtOMPDC showing the interactions of active site residues with BMP, an intermediate/transition state analog.

Also, at least two structural strategies can be envisioned for stabilization of the anionic intermediate as the transition state is formed: 1) attractive electrostatic

interactions with the ε-ammonium group of Lys 72 that protonates C6 of the intermediate (6); and 2) enhanced hydrogen-bonding interactions with the backbone amide group of Ser 127 (16, 17). For the latter strategy to be effective, the negative charge must be delocalized to O4 in the pyrimidine ring (as described by the carbene resonance structure; **Scheme 1**); calculations support the participation of this interaction, although its importance has not been tested directly (16, 17).

These destabilizing/stabilizing interactions are "enforced" by a closed, catalytically active conformation (E_c) that is observed crystallographically in the presence of the intermediate/transition state analogs 6-azaUMP (18) and BMP (11); in their absence, MtOMPDC exists in an open, catalytically inactive conformation (E_o) (**Figure 15**) (12, 13). The more favorable transition from E_o to E_c in the E_o•S and E_c•S complexes [K_c *vs.* K_c' in the kinetic model shown in **Scheme 2** and **Figure 16** (19); *vide infra*] is responsible, in part, for a 2.4×10^8-fold activation of the enzyme and involves 1) closure of the nine-residue active site loop (Pro 180 to Asp 188) that sequesters the substrate from solvent, and 2) formation of a more compact conformation for the $(\beta/\alpha)_8$-barrel structure that juxtaposes the substrate with the residues involved in substrate destabilization and intermediate stabilization.

Scheme 2



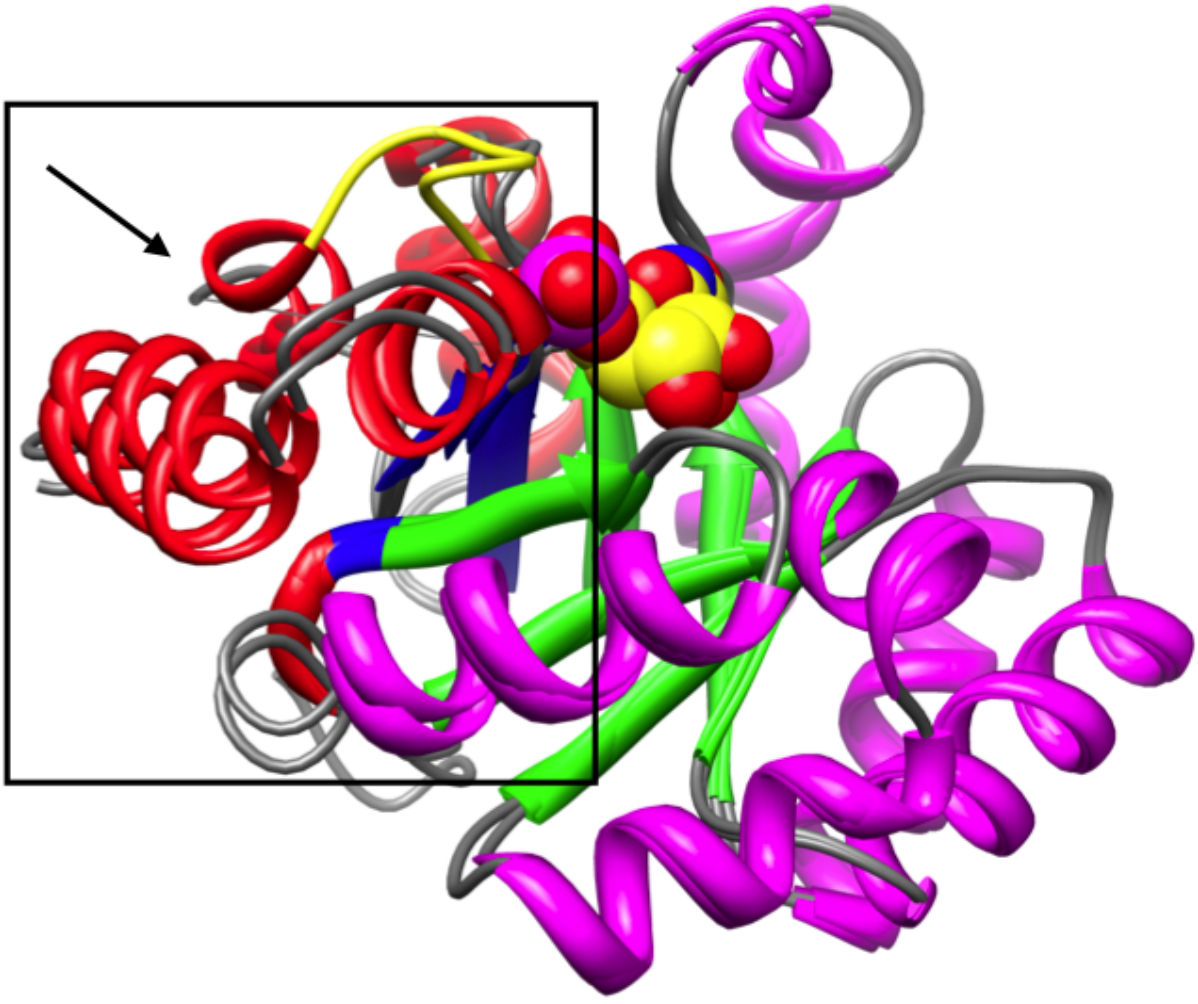


Figure 15. Superposition of the structures of individual polypeptides of unliganded (E_o) and BMP-liganded (E_c) MtOMPDC, showing the movement of the "mobile" domain (red a-helices and blue b-strands) toward the "fixed" domain (magenta a-helices and green b-strands). The ordered active site loop is colored yellow in the BMP-liganded structure.

Each $(\beta/\alpha)_8$-barrel in the obligate dimer is formed from two "half-barrel" domains, a "fixed" domain (including β-strands 2, 3, 4, and 5) located at the dimer interface and a "mobile" domain (including β -strands 6, 7, 8, and 1) (20); in the transition from E_o and E_c, the "mobile" domain moves toward the "fixed" domain, constricting the active site cavity that envelops the substrate (**Figure 2**).[2] Because the active site loop sequesters the substrate from solvent in the E_c•S complex, substrate binding to E_o must precede the conformational change. Thus, "induced fit", not "conformational selection", is responsible for altering the distribution between E_o and E_c when the substrate binds (21).[3]

The energy obtained from formation of the E_o•S complex from E_o + S is partitioned between 1) "gripping" the substrate [with the value of K_M providing a measure of the affinity for the substrate ($1/[K_c \bullet K_s]$ in **Scheme 2**; K_s is defined as an association complex but K_M is a dissociation constant)]; and 2) stabilizing both E_c•S and E_c•$S^{\ddagger}$ (*vide infra*) relative to E_o•S (K_c in **Scheme 2**). That E_c is unstable relative to E_o in the absence of substrate (K_c') facilitates release of the UMP product; if the total energy obtained from substrate binding were used only to "grip" the substrate, release of the product from the E_c•P complex likely would be rate-determining (22). Therefore, the structural and energetic coupling between substrate binding to E_o and the subsequent transition from E_o•S to E_c•S is essential for understanding the catalytic strategy that produces the extraordinary rate enhancement.

We recently reported initial efforts to identify structural features that stabilize E_c•S (19). Those focused on a hydrophobic cluster that includes Val 182 in the active site loop as well as Ile 199, Val 201, and Ile 218 located near the C-terminal ends of the seventh and eighth β-strands in the "mobile" domain. Substitutions of these residues that are "remote" from the binding site for the 5'-phosphate group with Ala increase the value of K_M but have little effect on the value of k_{cat}. The simplest interpretation is that the hydrophobic cluster stabilizes the closed conformation of the active site loop and, therefore, contributes to the stability of E_c•S. Without this stabilization, the formation of E_c•S from E_o + S is less favorable, requiring increased concentrations of substrate to provide the energy to enable its formation (as revealed by an increase in the value of K_M). However, when E_c•S is formed, it is fully competent to yield product via the vinyl anion/carbene intermediate and the transition state that precedes it on the reaction coordinate (E_c•$S^{‡}$) so the value of k_{cat} (the energy difference between E_c•S and E_c•$S^{‡}$) is unchanged.

In this article we focus on residues involved in hydrogen-bonding interactions between the "mobile" and "fixed" domains observed in E_c but not in E_o (involving Thr 159, Arg 160, and Tyr 206), also "remote" from the binding site for the 5'-phosphate group, to determine whether these contribute to the stability of E_c•S. We also target the only two residues (Gln 185 and Arg 203) that provide direct side chain contacts with the 5'-phosphate group of the substrate and, therefore, reasonably would be involved in generation of the intrinsic binding energy (IBE) of the 5'-phosphate group that is used to enable stabilization of both E_c•S and E_c•$S^{‡}$ relative to E_o + S (23). Finally, we

examine double mutants involving one "remote" residue that contributes to the stability of E_c•S (Val 182, Thr 159, or Arg 160) and one residue that generates the IBE of the 5'-phosphate group (Arg 203). The kinetic and structural properties of the mutants are consistent with a structure-based model in which E_c•S and E_c•$S^{‡}$ are stabilized equally relative to E_o•S by 1) the interaction of the 5'-phosphate group with its binding site that allows/directs reorganization of the active site loop, 2) assembly of the hydrophobic cluster near the N-terminus of the reorganized active site loop that stabilizes its closed conformation, and 3) formation of interactions across the domain interface in E_c•S and E_c•$S^{‡}$, some of which (those involving Thr 159 and Arg 160) are spatially proximal to the reorganized active site loop and, therefore, are influenced by its structure. This analysis clarifies the role of the IBE of the 5'-phosphate group in reducing the value of $\Delta G^{‡}$ and provides the framework for a more detailed investigation of the structural features involved in both generation of the IBE and its use to facilitate the conversion of E_o•S to both E_c•S and E_c•$S^{‡}$.

The conclusion that the IBE generated by interaction of the 5'-phosphate group with its binding site allows stabilization of *both* the E_c•S and E_c•$S^{‡}$ complexes, not just the E_c•$S^{‡}$ complex, relative to E_o + S by enabling formation of "remote" interactions that stabilize E_c likely can be applied to many (all?) enzymes in which conformational changes accompany substrate binding.

MATERIALS AND METHODS

Materials. OMP, EO , FEO, and BMP were prepared by published procedures (6, 11, 13, 24-26). All solutions were prepared with Millipore Ultrapure filtered water.

Protein Expression and Purification and Site-Directed Mutagenesis. The gene encoding MtOMPDC was previously cloned in the pET-15b vector (Novagen). The site-mutants were constructed (overlap extension) and the wild type and mutant proteins were expressed and purified following previously published procedures (13, 15, 24).

Assay for Decarboxylation of OMP. Assays were performed at 25 °C and in 10 mM MOPS, pH 7.1, containing 100 mM NaCl following published procedures (13, 15, 24).

Assay for Decarboxylation of EO. The values of k_{cat}/K_M for decarboxylation of EO were determined as previously described (15, 25).

Assay for Decarboxylation of FEO. For wild type MtOMPDC and single mutants, the values of k_{cat}/K_M for decarboxylation of FEO were determined by quantitating the first order decay of FEO at 290 nm using procedures developed for EO (15, 25).

For double and triple mutants, the values of k_{cat}/K_M for decarboxylation of FEO were determined using HPLC-based end point assays using procedures developed for EO (15, 25); the amount of FEU product was quantitated using 5-fluorouridine as the reference standard (ε = 9660 M^{-1} cm^{-1} in 0.1 N HCl).

Crystallization and Data Collection. Eight different crystal forms (**Table 2**) were grown by the sitting drop method at room temperature for single mutants of MtOMPDC

liganded with BMP: (1) Q185A•BMP, (2) R203A•BMP, (3) T159V•BMP, (4) T159A•BMP, (5) T159S•BMP, (6) R160A•BMP, (7) Y206F•BMP, and (8) K82A•BMP. The protein solutions for all eight crystallizations contained the corresponding protein (30 mg/mL) in 20 mM HEPES, pH 7.5, 100 mM NaCl, 3 mM DTT, and 40 mM BMP. The precipitants were the following:

(1) For the Q185•BMP complex, the precipitant contained 30% PEG 4000, 0.1 M Tris•HCl, pH 8.5, and 0.2 M lithium sulfate.

(2) For the R203A•BMP complex, the precipitant contained 2.4 M sodium malonate, pH 7.0.

(3) For the T159V•BMP complex, the precipitant contained 30% PEG 4000, 0.1 M sodium citrate, pH 5.6, and 0.2 M ammonium sulfate.

(4) For the T159A•BMP complex, the precipitant contained 20% PEG 4000, 20% 2-propanol, and 0.1 M sodium citrate, pH 5.6.

(5) For the T159S•BMP complex, the precipitant contained 20% PEG 4000, 20% 2-propanol, and 0.1 M sodium citrate, pH 5.6.

(6) For the R160A•BMP complex, the precipitant contained 30% PEG 4000, 0.1 M sodium citrate, pH 5.6, and 0.2 M ammonium acetate.

(7) For the Y206F•BMP complex, the precipitant contained 30% PEG 4000, 0.1 M HEPES, pH 7.5, and 0.2 M magnesium chloride.

(8) For the K82A•BMP complex, the precipitant contained 20% PEG 8000, 0.1 M Tris•HCl, pH 8.5, and 0.2 M magnesium chloride.

Ten different crystal forms (**Table 3**) were grown by sitting drop method at room temperature for double and triple mutants of MtOMPDC liganded with BMP: (1) Q185A•R203A•BMP, (2) T159V•V182A•BMP, (3) T159V•Y206F•BMP, (4) V182A•Y206F•BMP, (5) R160A•V182A•BMP, (6) R160A•Y206F•BMP, (7) R203A•V182A•BMP, (8) R203A•T159V•BMP, (9) R203A•R160A•BMP, and (10) T159V•V182A•Y206F•BMP. The protein solutions for all ten co-crystallizations contained the corresponding protein (30 mg/mL) in 20 mM HEPES, pH 7.5, 100 mM NaCl, 3 mM DTT, and 40 mM BMP. The precipitants were the following:

(1) For the Q185A•F203A•BMP complex, the precipitant contained 2.0 M ammonium sulfate, pH 4.9.

(2) For the T159V•V182A•BMP complex, the precipitant contained 0.4 M sodium phosphate, 1.6 M potassium phosphate, 0.1 M imidazole, pH 8.0, and 0.2 M sodium chloride.

(3) For the T159V•Y206F•BMP complex, the precipitant contained 1.4 M sodium citrate, and 0.1 M HEPES, pH 7.5.

(4) For the V182A•Y206F•BMP complex, the precipitant contained 0.8 M sodium phosphate, 1.2M potassium phosphate, and 0.1 M acetate, pH 4.5.

(5) For the R160A•V182A•BMP complex, the precipitant contained 30% PEG 4000, 0.1 M Tris•HCl, pH 8.5, and 0.2 M lithium sulfate.

(6) For the R160A•Y206F•BMP complex, the precipitant contained 30% PEG 4000, and 0.1 M Tris•HCl, pH 8.5, and 0.2 M magnesium chloride.

(7) For the R203A•V182A•BMP complex, the precipitant contained 1.0 M sodium citrate, and 0.1 M cacodylate, pH 6.5.

(8) For the R203A•T159V•BMP complex, the precipitant contained 0.8 M sodium phosphate, 1.2 M potassium phosphate, and 0.1 M acetate, pH 4.5.

(9) For the R203A•R160A•BMP complex, the precipitant contained 3.5 M sodium formate, pH 7.0.

(10) For the T159V•V182A•Y206F•BMP complex, the precipitant contained 60% tacsimate, pH 7.0.

Prior to data collection, the crystals of all 18 crystal forms (**Tables 2 and 3**) were transferred to cryoprotectant solutions composed of their mother liquids and 20% glycerol and flash-cooled in a nitrogen stream. The X-ray diffraction data sets (**Tables 2 and 3**) were collected at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector and at NSLS X29A beamline on the 315q CCD detector. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK (*27*). The data collection statistics are given in **Tables 2 and 3**

Table 2. Data collection and refinement statistics for BMP complexes of single mutants

	Q185A	R203A	T159V	T159A	T159S	R160A	Y206F	K82A
Data collection								
Space group	$P2_1$	$P2_12_12$	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
No. of molecules	2	2	2	2	2	2	2	2
in asym. unit								
Cell dimensions								
a (Å)	59.80	80.00	59.98	52.61	58.67	59.80	59.62	59.78
b (Å)	63.86	64.00	64.12	74.04	73.87	64.14	63.41	64.01
c (Å)	61.47	73.30	61.89	117.50	59.52	61.78	61.04	61.63
β (°)	115.28		115.46		119.50	115.67	115.02	115.52
Resolution (Å)	1.37	1.50	1.40	1.60	1.3	1.4	1.42	1.49
No. of unique reflections	87599	58327	79912	54292	104965	76701	76776	68389
R_{merge}	0.065	0.082	0.091	0.076	0.038	0.083	0.073	0.043
Completeness (%)	99.7	95.6	95.8	88.5	96.8	91.5	98.1	99.7
Refinement								
Resolution (Å)	25.0-1.37	25.0-1.50	25.0-1.40	25.0-1.60	25.0-1.3	25.0-1.40	25-1.42	25-1.49
R_{cryst}	0.165	0.223	0.188	0.204	0.165	0.199	0.172	0.153
R_{free}	0.178	0.247	0.206	0.231	0.176	0.222	0.193	0.174
No. atoms								
Protein	3421	3412	3364	3322	3418	3374	3379	3416
Waters	454	144	396	170	410	410	347	444
Bound ligands	BMP SO4	BMP	BMP	BMP GOL	BMP GOL, POL	BMP	BMP GOL	BMP GOL
Ligand atoms	54	44	44	50	82	44	50	50
R.m.s. deviations								
Bond lengths (Å)	0.006	0.005	0.006	0.007	0.006	0.006	0.006	0.006
Bond angles (°)	1.0	1.2	1.1	1.1	1.1	1.1	1.1	1.1
PDB entry	3V1P	3LI0	3P60	3P5Y	3P5Z	3P61	3RLV	3RLU

Table 3. Data collection and refinement statistics for BMP complexes of double and triple mutants

	Q185/ R203A	T159V/ V182A	T159V/ Y206F	V182A/ Y206F	R160A/ V182A	R160A/ Y206F	T159V/ V182A/ Y206F	R203A/ V182A	R203A/ T159V	R203A/ R160A
Data collection										
Space group	$P4_12_12$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_12_12_1$
No. of molecules in asym. unit	2	2	2	2	2	2	2	2	2	2
Cell dimensions										
a (Å)	91.51	59.77	59.93	59.56	59.77	60.08	59.81	55.08	54.35	56.62
b (Å)	91.51	64.07	63.67	63.62	64.07	63.83	63.98	66.40	63.06	56.61
c (Å)	135.38	61.64	61.43	61.26	61.64	61.91	61.43	59.55	54.51	127.21
β (°)		115.46	115.34	115.00	115.46	115.53	115.48	101.85	98.95	
Resolution (Å)	1.94	1.54	1.34	1.32	1.32	1.26	1.32	1.53	1.71	1.42
No. of unique reflections	41820	55241	87118	96995	95690	111994	97927	62002	38041	77257
R_{merge}	0.095	0.075	0.065	0.055	0.088	0.073	0.058	0.047	0.092	0.065
Completeness (%)	96.8	89.4	93.0	99.8	97.4	98.4	99.8	98.0	96.1	92.2
Refinement										
Resolution (Å)	25-1.94	25-1.54	25-1.34	25.0-1.32	25.0-1.32	25.0-1.26	25-1.32	25-1.53	25-1.71	25-1.4
R_{cryst}	0.172	0.173	0.173	0.165	0.231	0.176	0.164	0.146	0.225	0.184
R_{free}	0.198	0.206	0.192	0.178	0.247	0.196	0.183	0.172	0.268	0.211
No. atoms										
Protein	3254	3390	3477	3434	3327	3444	3426	3499	3314	3359
Waters		191	367	406	342	553	464	373	74	225
Bound ligands	BMP SO4	BMP GOL	BMP GOL	BMP GOL	BMP	BMP	BMP GOL	BMP GOL	BMP	BMP
Ligand atoms	54	50	50	50	44	44	50	69	44	44
R.m.s. deviations										
Bond lengths (Å)	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.008	0.006
Bond angles (°)	1.00	1.1	1.1	1.1	1.0	1.1	1.1	1.1	1.0	1.1
PDB entry	4FX8	3QEZ	3QF0	3QMT	3QMR	3SJ3	3QMS	4FX6	4FXR	4GC4

Structure Determination and Model Refinement. All 18 MtOMPDC structures (**Tables 2 and 3**) were solved by molecular replacement with fully automated molecular replacement pipeline BALBES (28), using only input diffraction and sequence data. Partially refined structures of all 18 MtOMPDC crystal forms (**Tables 2 and 3**) were the outputs from BALBES. Subsequent several iterative cycles of refinement were performed for each crystal form including: model rebuilding with COOT (29), refinement with PHENIX (30), and automatic model rebuilding with ARP (31).

All loops are well-defined in all 18 BMP liganded MtOMPDC structures. Usually some loops are disordered in unliganded MtOMPDC structures. All 18 MtOMPDC complexes have well-ordered density for all bound BMP ligands in every active site. No nonglycine residue lies in disallowed regions of the Ramachandran plot for all 18 MtOMPDC structures shown in **Tables 2 and 3**. All 18 structures form stable dimers. The monomers within each dimer are connected by a noncrystallographic two-fold axis in all 18 structures.

Representative electron density for the Q185A mutant liganded with BMP is shown in **Figure 16**.

Final crystallographic refinement statistics for all determined MtOMPDC structures are provided in the **Tables 2 and 3**.

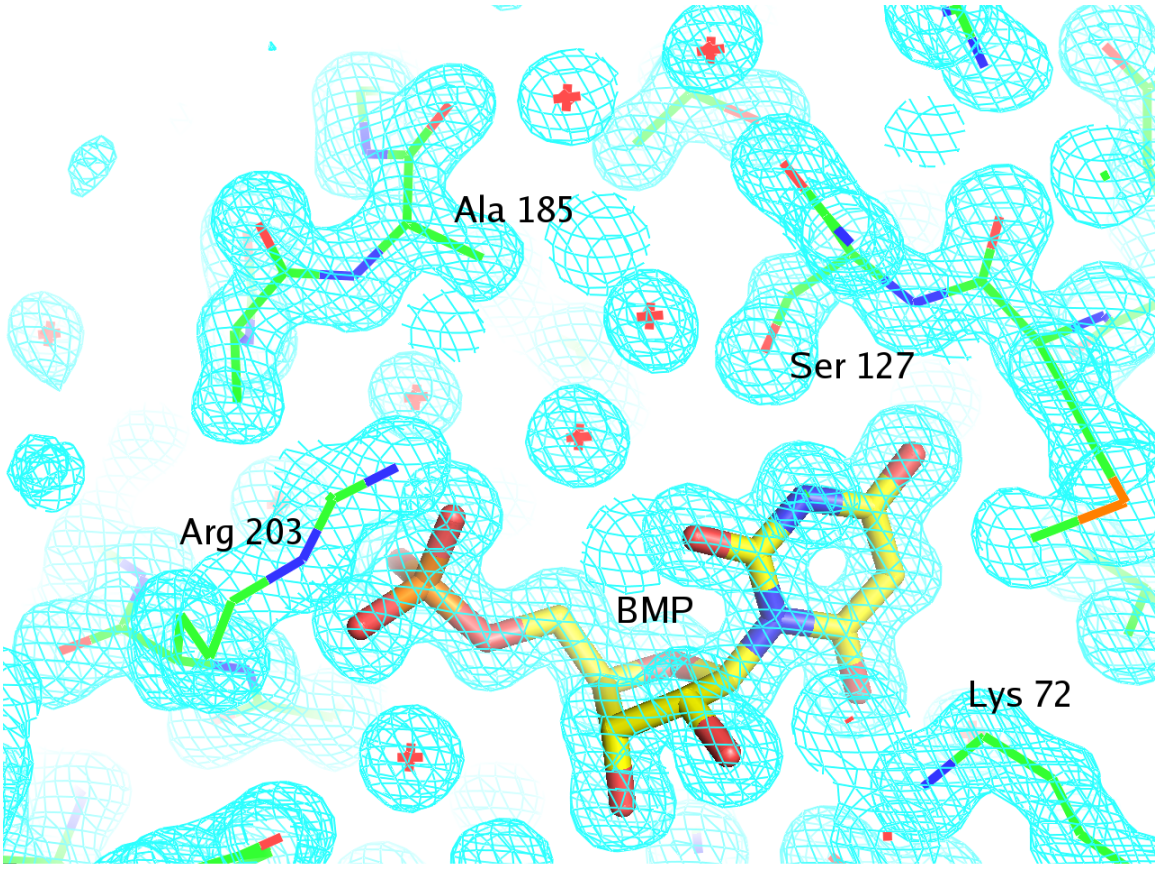


Figure 16. Representative electron density map for the active site of the Q185A mutant complexed with BMP and contoured at 1.5 Å. The figure was produced with PyMOL (41). The details of the interactions between BMP and the active site are described in the text.

RESULTS AND DISCUSSION

Intrinsic Binding Energy (IBE) of the 5'-Phosphate Group. As described by Jencks (23), the IBE of a remote substituent, e.g., the 5'-phosphate group of OMP, is used "to pay for substrate destabilization through distortion, electrostatic interactions, and desolvation, for bringing about the necessary loss of entropy by freezing the substrates in the proper position for reaction, and for binding to the transition state. The maximum binding energy is then not realized directly in the binding of the substrate, but is more completely realized in the transition state." For MtOMPDC, the 5'-phosphate group of OMP "activates" decarboxylation 2.4 x 10^8-fold as measured by comparing the values of k_{cat}/K_M for OMP and EO that lacks the 5'-phosphate group [**Scheme 2** (19, 26, 32)], 2.9 x 10^6 M^{-1} sec^{-1} and 12.4 x 10^{-3} M^{-1} sec^{-1}, respectively (**Table 4**); the IBE calculated from this ratio is 11.4 kcal/mol (**Table 4**)].

The challenge has been to provide a structure-based understanding for how the IBE of a "remote" group of the substrate is both generated and used to activate the enzyme. Given the conformational change induced by ligand binding (**Figure 15**), a plausible explanation is that in MtOMPDC the IBE of the 5'-phosphate group not only increases the affinity of the enzyme for the substrate [K_s (and K_s') in the kinetic model in **Scheme 2**/**Figure 17**; green arrows in **Figure 17**] but also increases the fraction of the enzyme that exists as E_c•S instead of E_o•S ($K_c >> 1$ in contrast to $K_c' << 1$ for E_c and E_o; red and blue arrows, respectively, in **Figure 17**) by enabling stabilizing contacts between the "mobile" and "fixed" domains.

Table 4. Kinetic Constants for OMP at pH 7.1 and 25 °C.

MtOMPDC	k_{cat} (s^{-1})	K_m (mM^{-1})	k_{cat}/K_m (M^{-1} s^{-1})	Fold change k_{cat}	Fold change K_m	Fold change k_{cat}/K_m	$\Delta\Delta G^{\ddagger}$ (kcal/mol)[b]
Wild type	5.3 ± 0.1	1.8 ± 0.5	2.9×10^6	-	-	-	-
T159V	2.3 ± 0.1	33 ± 2	7.0×10^4	2	18	41	2.2
R160A	2.0 ± 0.1	30 ± 1	6.7×10^4	3	17	43	2.2
V182A	3.0 ± 0.1	24 ± 1	1.3×10^5	2	13	22	1.8
Y206F	4.1 ± 0.2	5.5 ± 1.0	7.5×10^5	1	3	4	0.8
T159V/V182A	0.56 ± 0.02	700 ± 90	8.0×10^2	10	390	3600	4.8
T159V/Y206F	1.1 ± 0.1	300 ± 50	3.7×10^3	5	170	780	3.9
R160A/V182A	1.0 ± 0.1	600 ± 60	1.7×10^3	5	330	1700	4.4
R160A/Y206F	1.1 ± 0.1	130 ± 10	8.5×10^3	5	72	340	3.4
V182A/Y206F	1.5 ± 0.1	100 ± 15	1.5×10^4	4	56	190	3.1
T159V/V182A/Y206F	0.24 ± 0.01	1900 ±	1.3×10^2	22	1100	2.2×10^4	5.9
Q185A	1.4 ± 0.1	110 ± 15	1.3×10^4	4	61	220	3.1
R203A	1.5 ± 0.1	980 ± 10	1.5×10^3	4	540	1900	4.4
Q185A/R203A	-	-	9.6[a]	-	-	3.0×10^5	7.4
T159V/R203A	-	-	9.9[a]	-	-	3.0×10^5	7.4
R160A/R203A	-	-	16[a]	-	-	1.8×10^5	7.1
V182A/R203A	-	-	32[a]	-	-	9.1×10^4	6.7

[a]Velocity is directly proportional to substrate concentration.
[b]For k_{cat}/K_M.

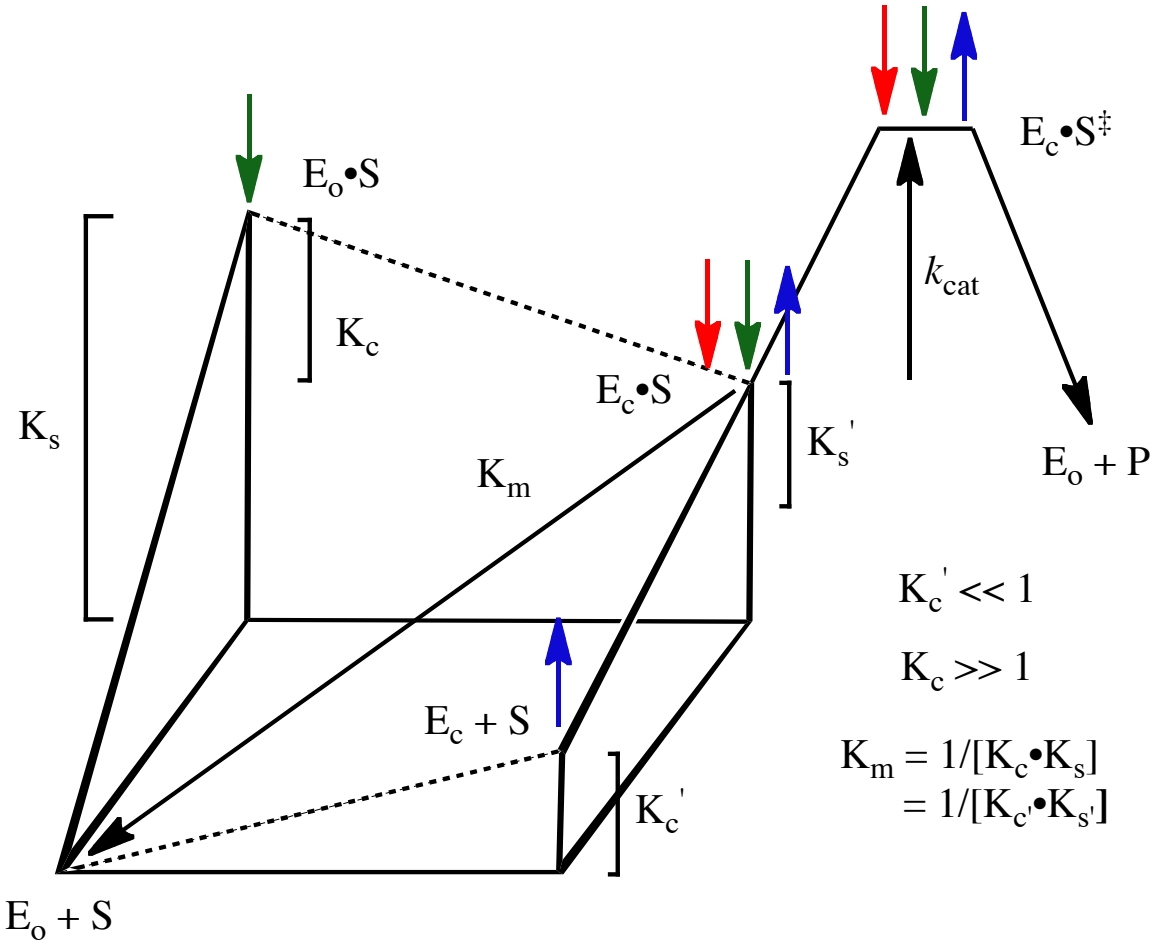
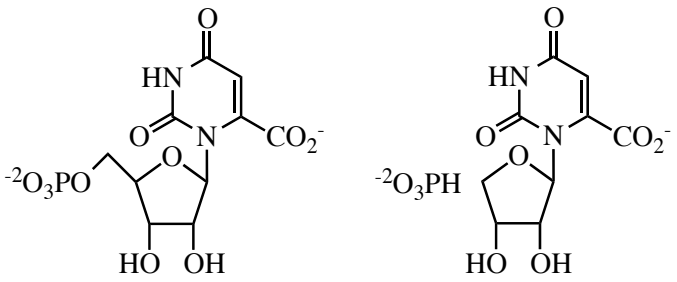


Figure 17. Three-dimensional free energy diagram for the kinetic model depicted in Scheme 2. See the text for a detailed explanation.

Scheme 3



OMP phosphite + EO

Model to Explain the Role of the IBE of the 5'-Phosphate Group in the Activation of MtOMPDC. We previously proposed the kinetic model in **Scheme 2** to understand the relationships among E_o, E_c, S, and the values of K_M and k_{cat} (19); the three-dimensional energy diagram for the model is shown in **Figure 17**.

In the model, E_o predominates in the absence of substrate ($K_C' << 1$); E_c predominates in its presence ($K_C >> 1$; as E_c•S). The values of K_c' and K_c are determined by the relative stabilities of E_o and E_c in the absence and presence of substrate, respectively. The dissociation constant for S, $[E_o][S]/[E_c•S]$, is assumed to be the value of K_M, $1/[K_c•K_s]$ ($= 1/[K_c'•K_s']$).[3]

The value of the IBE for the 5'-phosphate group is obtained from the increase in the value of k_{cat}/K_M associated with the 5'-phosphate group occupying its binding site in E_c (from the ratio of the values for OMP and EO). Because the carboxylate group of the substrate (in either OMP or EO) is distal from the binding site for the 5'-phosphate group (**Figure 14**), the value of k_{cat} (the energy difference between E_c•S and E_c•$S^{‡}$) can be expected to be independent of the absence/presence of the 5'-phosphate group in the absence of "long range" interactions with the carboxylate group of the substrate and/or the anionic intermediate, i.e., the same value for OMP and EO *if it were possible to fully populate E_c•S when EO is the substrate*. The model also assumes that interaction of the 5'-phosphate group is required to stabilize E_c•S relative to E_o•S (K_c *vs.* K_c' in the unliganded enzyme), so the value of k_{cat} for EO is determined by the fraction of the unliganded enzyme that exists as E_c, i.e., from the value of K_c'.

The energy obtained from the 5'-phosphate group interacting with its binding site (available to OMP but not EO) can be used to 1) increase the stabilities of the E_o•S and E_c•S complexes relative to E_o + S (green arrows in **Figure 17**; "gripping" the substrate); and 2) increase the stability of E_c•S relative to E_oS (red arrows in **Figure 17**; stabilizing the active conformation). Both effects stabilize the E_c•S complex relative to E_o + S, thereby decreasing the value of K_M and increasing the value of k_{cat}/K_M. Thus, the activation by the 5'-phosphate group is caused by two independent effects that cannot be deconvoluted: effects on K_s and/or K_c when OMP is the substrate. If the value of k_{cat} is independent of the interactions of the 5'-phosphate group with its binding site (evaluated by the loss of individual interactions by site-directed mutagenesis; *vide infra*), the activation by the 5'-phosphate group is the result of 1) an increase in the affinity of the enzyme for the OMP substrate (reflected in the value of K_s), and 2) an increase in the fraction of the enzyme that exists as E_c (as E_c•S, i.e., K_c *vs*. K_c').

The stability of E_c•S relative to E_o + S (measured by the K_M) is also influenced by residues "remote" from the binding site for the 5'-phosphate group. The structures of the unliganded and BMP-liganded structures of MtOMPDC allow the proposal that the hydrophobic cluster involving Val 182 in the active site loop as well as hydrogen bonds involving Thr 159, Arg 160, and Tyr 206 between the "mobile" and "fixed" domains observed in E_c but not in E_o contribute to the stability of E_c relative to E_o both in the absence and presence of ligand (K_c' and K_c, respectively; blue arrows in **Figure 17**).

In our initial studies of this model, we observed that Ala substitutions for residues in the hydrophobic cluster (Val 182, Ile 199, Val 201, and Ile 218) that are "remote" from the 5'-phosphate group and its binding site increase the value of K_M but do not influence the value of k_{cat} (19). We determined the structures of the mutants in the presence of BMP and observed that they are superimposable on that of wild type with the exception of the side chains of the substituted residues. Thus, the changes in the values of K_M could be associated with an incomplete "remote" hydrophobic cluster that reduces the values of both K_c' and K_c (blue arrows in **Figure 17**); the unchanged values for k_{cat} were explained by the unperturbed interactions of the active site with the OMP substrate and transition state in the $E_c \bullet S$ and $E_c \bullet S^{\ddagger}$ complexes (**Figure 14**), so the energy difference between $E_c \bullet S$ and $E_c \bullet S^{\ddagger}$ is unchanged.

The Ala substitutions in the hydrophobic core do not alter the values of the IBEs, i.e., the values of k_{cat}/K_M for OMP and EO are equivalently reduced by the substitutions (19). Because the values for EO measure the fraction of the unliganded enzyme present as E_c (the value of K_c'), the equivalent reductions demonstrate that the interaction of the 5'-phosphate group with its binding site is not the sole determinant of the stability of the $E_o \bullet S$ complex relative to $E_o + S$ ($K_M = 1/[K_c \bullet K_s]$) and of E_c relative to E_o in the $E_c \bullet S$ and $E_o \bullet S$ complexes (K_c).

In the remainder of this manuscript and in the context of this model, we investigate the roles of the hydrogen-bonding interactions involving Thr 159, Arg 160, and Tyr 206 between the "mobile" and "fixed" domains observed in E_c but not in E_o as

well as Gln 185 and Arg 203 that contribute to the binding site for the 5'-phosphate group in the activation by the 5'-phosphate group of OMP.

Identification of Interactions at the Domain Interface That Stabilize E_c. MtOMPDC undergoes a conformational change from E_o in the absence of a ligand to E_c in the presence of BMP. As discussed in the Introduction, this conformational change involves closure of the active site loop (Pro 180 to Asp 188; yellow backbone in **Figure 15**) as well as the movement of a "mobile" domain (β-strands 6, 7, 8, and 1; red α-helices and blue β-strands in polypeptide A in **Figure 18**) toward a "fixed" domain (β-strands 2, 3, 4, and 5; magenta α-helices and green β-strands), with the "fixed" domains forming the dimer interface.

When the active site loop becomes ordered in E_c (yellow backbone in **Figures 15 and 16**), the hydrophobic cluster involving Val 182 in the active site loop is assembled [Val 182, Ile 199, Val 201, and Ile 218 (19); cyan residues in **Figure 18**], and hydrogen-bonding contacts also "remote" from the binding site for the 5'-phosphate group are established between the "mobile" and "fixed" domains (Tyr 206 with Lys 82 in the symmetry related polypeptide in the dimer, Thr 159 with the backbone carbonyl oxygen of Ser 127, and Arg 160 with the backbone carbonyl oxygens of both Met 126 and His 128; dark cyan residues in **Figure 18**). Enlarged views of these interactions are shown in **Figure 19** (Gln 185 and Arg 203 that interact directly with the 5'-phosphate group are colored brown). Our hypothesis is that these "remote" interactions at the domain interface also contribute to the stability of E_c relative to E_o as we previously concluded for the hydrophobic cluster (19).

Thr 159, Arg 160, and Tyr 206 in the "mobile" domain were targeted, with the expectation that substitutions of these with Ala (Thr 159 and Arg 160) or Phe (Tyr 206) would destabilize E_c (as determined by the values of k_{cat}/K_M for EO). The substitutions were not expected to alter the IBE, as previously observed for the residues that participate in the hydrophobic cluster at the base of the active site loop (19), because they are "remote" from the binding site for the 5'-phosphate group.

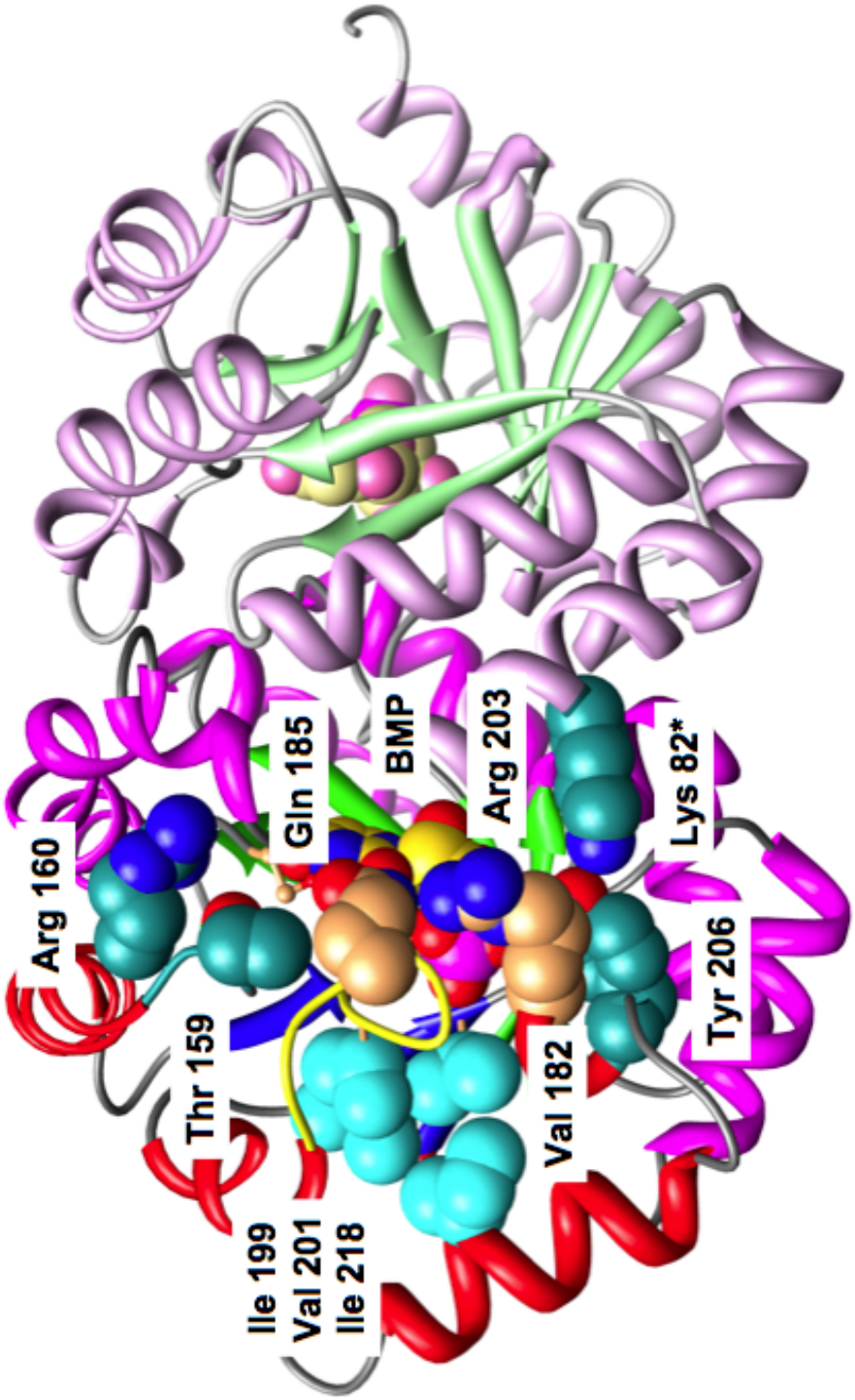


Figure 18. The structure of the dimer of MtOMPDC liganded with BMP, with the residues studied in this manuscript labeled. Ile 199, Val 201, and Ile 218 that participate in the hydrophobic cluster near the N-terminus of the active site loop (with Val 182) are also labeled (19).

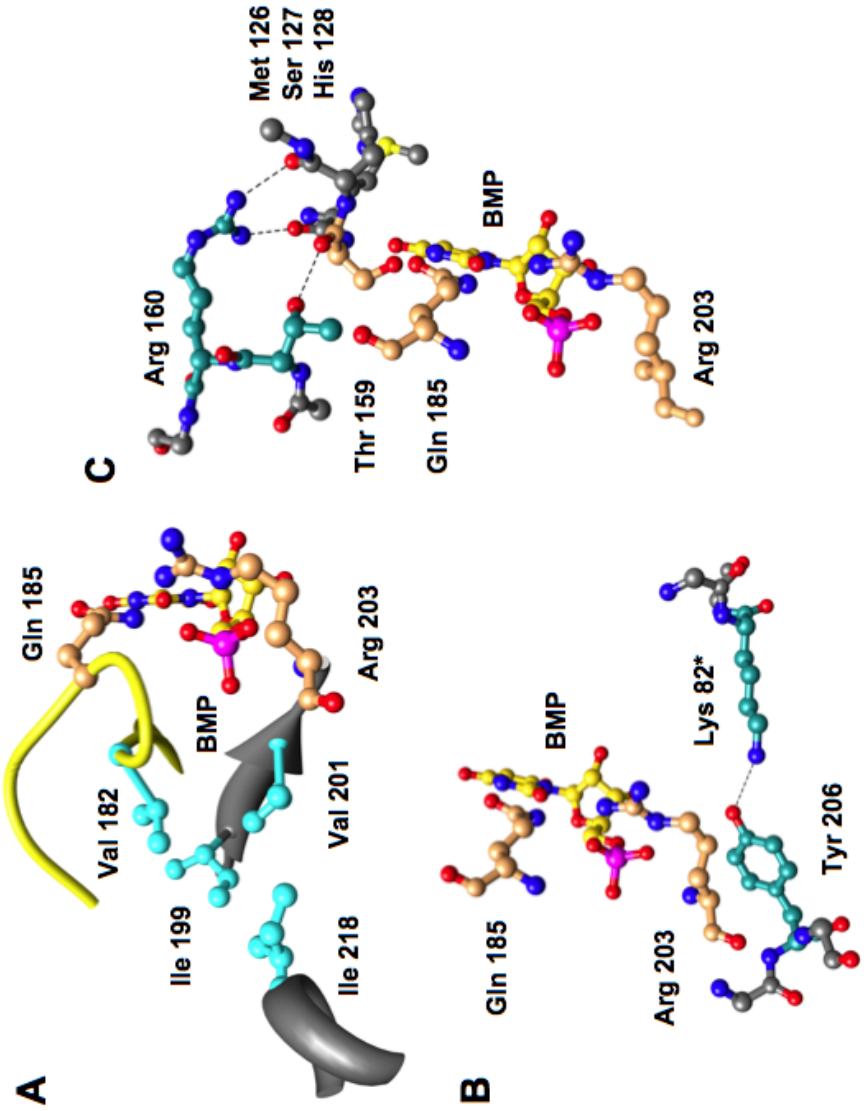


Figure 19. Interactions formed by closure of the "mobile" and "fixed" domains. Panel A, the hydrophobic cluster involving Val 182 at the N-terminus of the active site loop. Panel B, the hydrogen bond between Tyr 206 and Lys 82* in the symmetry-related polypeptide in the dimer. Panel C, the hydrogen bonds involving 1) the OH group of Thr 159 and the backbone carbonyl oxygen of Ser 127 and 2) the guanidinium group of Arg 160 and the backbone carbonyl oxygens of Met 126 and His 128.

Identification of Interactions That Generate the IBE of the 5'-Phosphate Group. Each of the nonesterified oxygens of the 5'-phosphate group participates in three hydrogen-bonds (a total of nine). As shown in **Figure 20**, panel A, the 5'-phosphate group of BMP directly contacts the carboxamide nitrogen of the side chain of Gln 185 in the active site loop (via a single hydrogen bond) and the guanidinium group of Arg 203 at the end of the eighth b-strand in the "mobile" domain (via two hydrogen bonds); it also contacts the backbone amide groups of Gly 202 and Arg 203 (via a single hydrogen bond to each). The importance of the interactions with the side chains of Gln 185 and Arg 203 in determining the activation by the 5'-phosphate group can be quantitated by characterization of mutants in which the side chains are "removed", e.g., the Q185A and R203A substitutions. However, the importance of the interactions with the backbone amide groups cannot be investigated by site-directed mutagenesis-based approaches.

In addition, the 5'-phosphate group is hydrogen-bonded to four water molecules, with three participating in hydrogen bonds to the backbone amide groups of Gly 181, Val 182, and Gln 185 in the active site loop; the fourth water molecule is hydrogen-bonded to both Tyr 206 in the "mobile" domain and Asp 20 in the "fixed" domain. The importance of these "indirect" contacts of the 5'-phosphate group with the protein also is difficult to investigate using site-directed mutagenesis-based approaches.

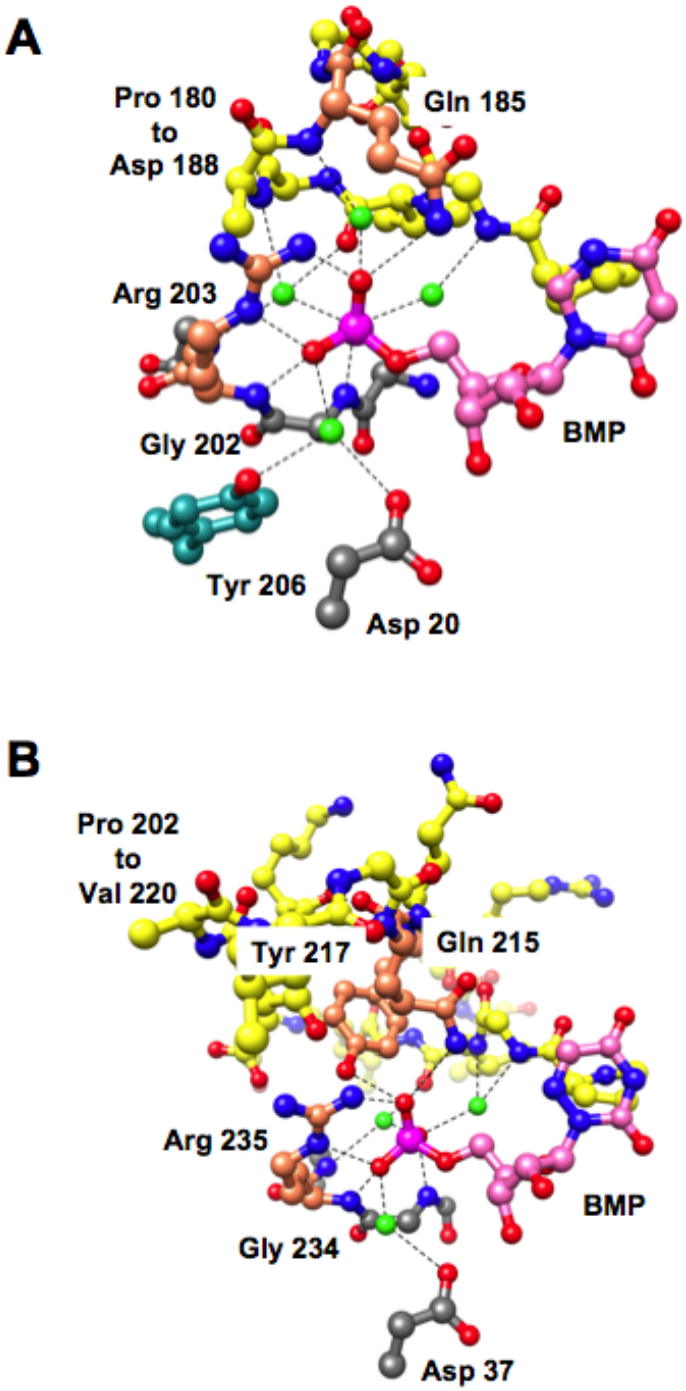


Figure 20. Hydrogen-bonding interactions of the nonesterified oxygens of the 5′-phosphate group of BMP. Panel A, the active site of MtOMPDC (details are provided in the text). Panel B, the active site of ScOMPDC (details are provided in the text). The water molecules that are involved in hydrogen bonds to the 5′-phosphate groups are colored green.

Our hypothesis is that *all* of the contacts made by the 5'-phosphate group are used to generate the total interaction energy that is partitioned between 1) "gripping" the substrate (K_s in **Scheme 2**/**Figure 17**; green arrows); and 2) stabilizing both E_c•S and E_c•$S^{\ddagger}$ (*vide infra*) relative to E_o•S (K_c in **Scheme 2**/**Figure 17**; red arrows). When the active site loop assumes its closed conformation when Gln 185 interacts with the 5'-phosphate group and the hydrophobic cluster involving Val 182 is formed, hydrogen-bonding contacts involving the loop's backbone, including the carbonyl oxygen of Pro 180, reposition the spatially proximal Thr 159 and Arg 160 in the "mobile" domain so that their side chains can form hydrogen bonds across the domain interface to the backbone carbonyl oxygens of Ser 127 as well as those of Met 126 and His 128 that flank Ser 127 in the "fixed" domain. The carboxamide oxygen of Gln 185 also forms a hydrogen bond with the OH group of Ser 127 in the "fixed" domain. Thus, we propose that IBE generated by the 5'-phosphate group is used, in part, to enable assembly of the hydrophobic cluster at the base of the active site loop as well as the interactions across the domain interface, with these providing independent contributions to the stabilization of E_c•S and E_c•$S^{\ddagger}$ relative to E_o•S (**Figures 18 and 19**).

Gln 185 and Arg 203 were targeted for site-directed mutagenesis, with the expectation that Ala substitutions would destabilize E_c•S and E_c•$S^{\ddagger}$ relative to E_o + S as judged by a decrease in the IBE as the result of a reduction in the value of k_{cat}/K_M for OMP but not for EO. Although the Q185A could be expected to decrease the value of K_c because of the interaction of the carboxamide group with Ser 127 in the "mobile"

domain (**Figure 14**; as measured by the value of k_{cat}/K_M for EO), R203A is not expected to alter the value of K_c because it is "remote" from the domain interface.

Structures of Mutants of Residues that Stabilize E_c. Five single mutants of residues that stabilize E_c [T159A, T159S, T159V, R160A, and Y206F; V182A was previously described and structurally characterized (19)] as well as six double/triple mutants that combine four of these (T159V/V182A, T159V/Y206F, R160A/V182A, R160A/Y206F, V182A/Y206F, and T159V/V182A/Y206F) were constructed and subjected to structural characterization in the presence of BMP. In these structures, the catalytic triads, Asp 70, Lys 72, and Asp 75* (from the symmetry related polypeptide in the dimer) as well as Ser 127 that hydrogen-bonds to the pyrimidine ring, superimpose well on the wild type enzyme (**Figure 21**). With the exception of the T159A substitution, the only structural changes are those associated with the side chain of the substituted residue; in the case of T159A, the conformation of the loop flanking the substitution (Pro 157 – Arg 163) is altered. Not surprisingly, the T159A mutant displayed significantly compromised kinetic constants (the value of k_{cat}/K_M is reduced 10^4-fold from that measured for wild type; data not shown) compared to all of the other mutants of residues that stabilize E_c. Our observation that the values of k_{cat} measured for the remaining mutants are "unchanged" relative to that measured for the wild type enzyme is consistent with the observation that the structures of the active sites are not perturbed by the substitutions.

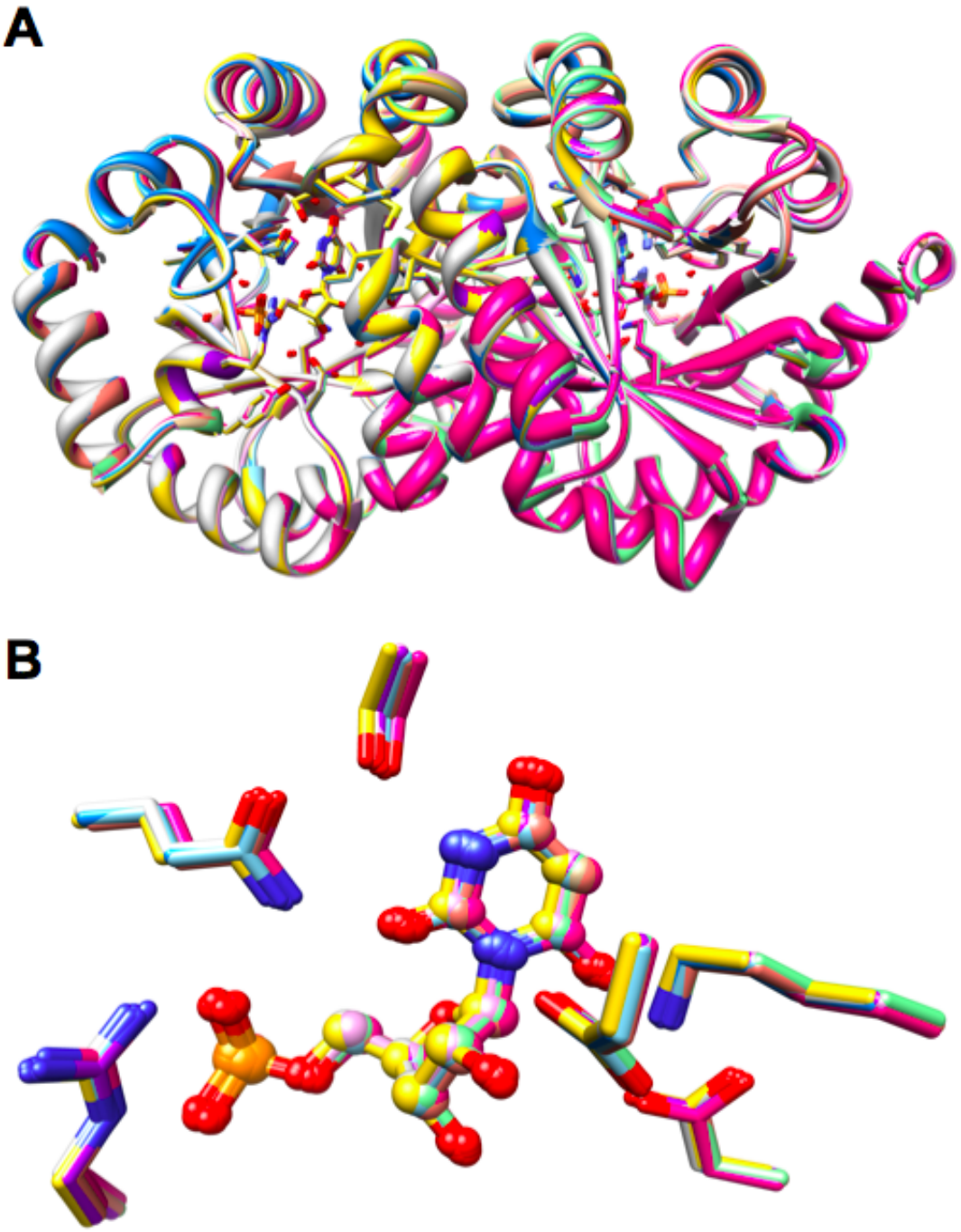


Figure 21. Superposition of the BMP-liganded structures of wild type MtOMPDC and the K82A, T159V, R160A, Y206F, T159V/V182A, T159V/Y206F, V182A/Y206F, and T159V/V182A/Y206F mutants of residues that stabilize E_c. Panel A, the polypeptide dimers. Panel B, structures of the active sites.

Structures of Mutants of Residues that Generate the IBE of the 5'-Phosphate Group. We constructed the Q185A and R203 single substitutions as well as the Q185A/R203A double substitution to remove the direct contacts between the 5'-phosphate group and the side chains of the protein. In the structure of the Q185A/R203A double mutant, the active site loop (Pro 180 to Asp 188) is partially disordered in the A polypeptide (Ala 184 and Ala 185; with increased B-values for the remaining "visible" loop residues) and assumes an altered conformation (also with increased B-values) in the B polypeptide (**Figure 22**).

In the mutant proteins, the structures of the active sites are not affected by the substitution(s); this is consistent with our observation that the values of k_{cat} for the Q185A and R203A single mutants are "unchanged" from that measured for wild type enzyme (*vide infra*); the value of k_{cat} could not be measured for the Q185A/R203A double mutant because the value of K_M is significantly elevated (K_s is significant decreased) by the loss of these contacts with the 5'-phosphate group. In the Q185A and Q185A/R203A mutants, two "new" water molecules occupy the position of the carboxamide group of Gln 185 in the wild type enzyme: one is located in the same position as the carboxamide amino group and is hydrogen-bonded to one of the nonesterified phosphoryl oxygens; the second is located in the same position as the carboxamide carbonyl oxygen and is hydrogen-bonded to both the first water molecule and the OH group of Ser 127. This hydrogen-bonded "chain" maintains an indirect interaction of the 5'-phosphate group with the "fixed" domain and, therefore, is expected to contribute the stability of the domain interface although the carboxamide

group is "missing". Finally, in the R203A and Q185A/R203A mutants, a "new" water molecule located in the same position as the e-nitrogen of the guanidinium group in the wild type protein is hydrogen-bonded to a second nonesterified oxygen.

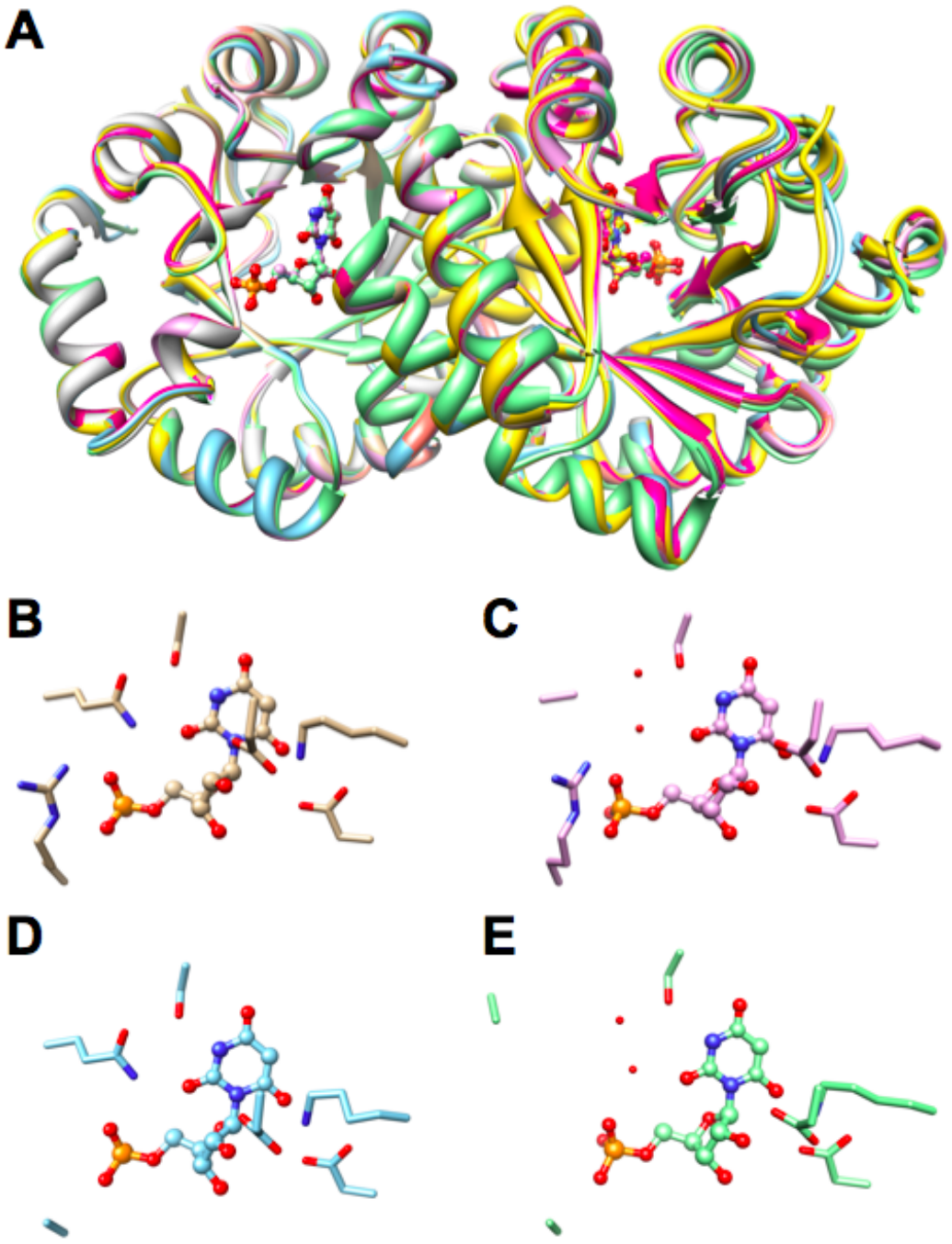


Figure 22. Superposition of the BMP-liganded structures of wild type MtOMPDC and the Q185A, R203A, and Q185A/R203A mutants of residues that generate the IBE of the 5'-phosphate group. Panel A, the polypeptide dimers. Panel B, active site of wild type. Panel C, active site of the Q185A mutant. Panel D, Active site of the R203A mutant. Panel E, active site of the Q185A/R203A mutant. The red spheres represent water molecules.

Structures of Double Mutants of Residues that Stabilize E_c and Generate the IBE of the 5'-Phosphate Group. The R203A substitution was combined separately with the spatially remote T159V, R160A, and V182A substitutions. As expected based on the structures of the single substitutions and the unchanged values for k_{cat} (*vide infra*), the structures of the active sites were unchanged in the mutants (**Figure 23**). Also, as in the structure of R203A, in each structure a "new" water molecule is located in the same position as the e-nitrogen of the guanidinium group in the wild type protein.

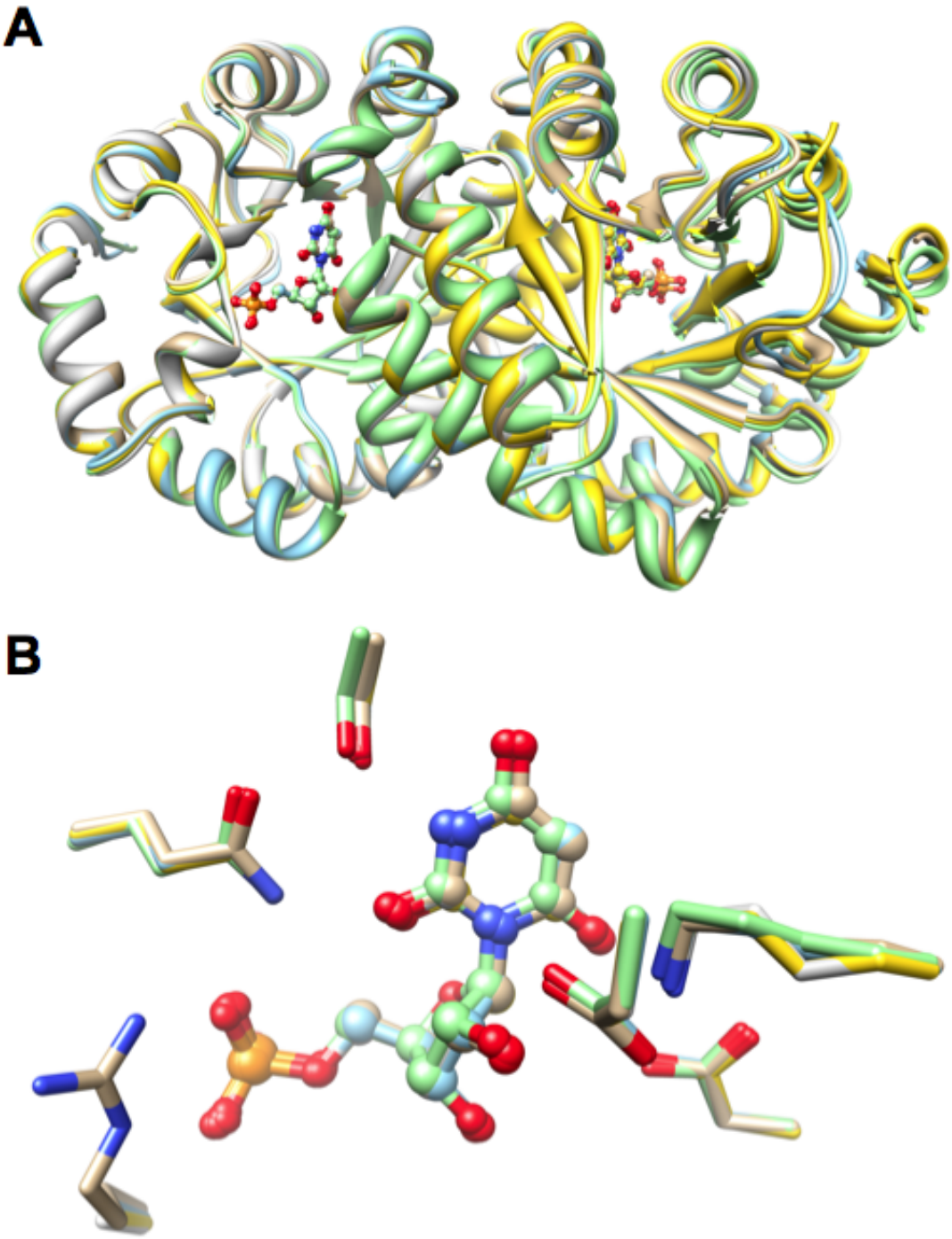


Figure 23. Superposition of the BMP-liganded structures of wild type MtOMPDC and the T159V/R203A, R160A/R203A, and V182A/R203A double mutants of residues that stabilize E_c and generate the IBE of the 5'-phosphate group. Panel A, the polypeptide dimers. Panel B, structures of the active sites.

Kinetic Constants for Mutants of Residues that Stabilize E_c. The values of k_{cat}, K_M, and k_{cat}/K_M using OMP as substrate were determined for the wild type and mutants that contain "remote" substitutions for residues hypothesized to stabilize E_c relative to E_o (**Table 4**). The values of k_{cat}/K_M for EO and/or FEO, $(k_{cat}/K_M)/K_D$ for activation of decarboxylation of EO by phosphite (the ability of phosphite to mimic the 5'-phosphate group), and the IBEs for the 5'-phosphate group also were determined (**Table 5**).

For the double/triple mutants, the more reactive FEO was used to estimate the values of k_{cat}/K_M for EO (*25, 33, 34*). For these mutants, the values of K_M for OMP and, therefore, both EO and FEO are significantly elevated because the fraction of the enzyme that exists as E_c is reduced (K_c' is reduced; blue arrows in **Figure 17**) because interactions that stabilize E_c are absent. Because of the increased reactivity of FEO, lower concentrations of enzyme (and/or shorter reaction times) can be used to measure the values of k_{cat}/K_M, thereby making these measurements possible. The correction factor for the double/triple mutants, 550, was obtained by averaging the values of the ratios of k_{cat}/K_M for EO and FEO for wild type MtOMPDC and the single mutants (**Table 5**).

For the single and most of the double mutants using OMP as substrate, the values of k_{cat} are "unchanged" relative to that measured for wild type enzyme; in contrast, the values of K_M are significantly increased. [The values for k_{cat} for OMP are (modestly) decreased for the T159V/V182A and

Table 5. Kinetic Constants for EO and FEO and IBEs of the 5'-Phosphate Group of OMP at pH 7.1 and 25 °C.

MtOMPDC	k_{cat}/K_m OMP[a] (M^{-1} s^{-1})	$\Delta\Delta G^{\ddagger}$ (kcal/mol)	k_{cat}/K_m EO[b] (M^{-1} s^{-1})	$\Delta\Delta G^{\ddagger}$ (kcal/mol)	k_{cat}/K_m FEO[b] (M^{-1} s^{-1})	FEO/EO	$(k_{cat}/K_m)/K_D$ EO•HP_i[b] M^{-2} s^{-1}	$\Delta\Delta G^{\ddagger}$ (kcal/mol)	5'-Phosphate IBE[c] kcal/mol
Wild type	2.9×10^{6}	-	12.4×10^{-3}	-	4.8	390	2300	-	11.4[d]
T159V	7.0×10^{4}	2.2	0.4×10^{-3}	2.0	0.083	210	35	2.4	11.2
R160A	6.7×10^{4}	2.2	0.2×10^{-3}	2.4	0.21	1100	56	2.1	11.6
V182A	1.3×10^{5}	1.8	1.4×10^{-3}	1.3	0.63	470	150	1.6	10.8
Y206F	7.5×10^{5}	0.8	6.0×10^{-3}	0.4	2.2	370	280	1.2	11.0
T159V/V182A	8.0×10^{2}	4.8	6.2×10^{-5e}	3.1	0.034				9.6
T159V/Y206F	3.7×10^{3}	3.9	6.3×10^{-5e}	3.1	0.035				10.5
R160A/V182A	1.7×10^{3}	4.4	4.6×10^{-5e}	3.3	0.025				10.3
R160A/Y206F	8.5×10^{3}	3.4	5.8×10^{-5e}	3.1	0.032				11.1
V182A/Y206F	1.5×10^{4}	3.1	5.5×10^{-4e}	1.8	0.3				10.1
T159V/V182A/Y206F	1.3×10^{2}	5.9	3.1×10^{-5e}	3.5	0.017				9.0
Q185A	1.3×10^{4}	3.1	2×10^{-3}	1.0	0.92	470	24	2.7	9.2
R203A	1.5×10^{3}	4.4	5.5×10^{-3}	0.4	4.7	860	2.4	4.0	7.4
Q185A/R203A	9.6	7.4	3.2×10^{-3}	0.8					4.7
T159V/R203A	9.9	7.4	0.2×10^{-3}	2.4					6.4
R160A/R203A	16	7.1	0.2×10^{-3}	2.4					6.6
V182A/R203A	32	6.7	1.3×10^{-3}	1.3					5.9

[a]From Table 3.
[b]The estimated error is ± 10%.
[c]Calculated from the ratio of the values of k_{cat}/K_M for OMP and EO.
[d]A lower estimate because the value of k_{cat}/K_M for OMP is partially diffusion-controlled (*24*).
[e]Calculated from the value of k_{cat}/K_M for FEO by dividing the value by 550 (the average ratio of the values for k_{cat}/K_M for EO and FEO for the wild type and those mutants for which the value for EO could be reliably measured).

T159V/V182A/Y206F mutants, suggesting small changes in structure and/or dynamics.] These observations are similar to those we reported for Ala substitutions for the residues in the hydrophobic cluster at the base of the active site loop that includes Val 182 (19). The increases in the values of K_M for the double/triple mutants are approximately multiplicative relative to the single mutants (additive in free energy), providing support that these "remote" residues provide "independent" contributions that stabilize E_c•S relative to E_o + S.

The values of k_{cat}/K_M for EO for the single mutants are decreased from that measured for wild type; the calculated values for the double/triple mutants are more significantly reduced, with the changes again approximately multiplicative relative to the single mutants. That the fold-decreases in the values of k_{cat}/K_M for OMP and EO are equivalent reinforces the interpretation that the targeted residues are important in determining the stability of E_c, i.e., the substitutions give decreased values for K_c and K_c' (blue arrows in **Figure 17**).

The values of $(k_{cat}/K_M)/K_D$ for activation of the decarboxylation of EO by phosphite were determined for the single mutants; these demonstrate that phosphite retains the ability to activate the mutants. That the values of k_{cat}/K_M for OMP, k_{cat}/K_M for EO, and $(k_{cat}/K_M)/K_D$ for phosphite are equivalently decreased is consistent with the interpretation that the changes are determined by the values of K_c, i.e., the 5'-phosphate group of OMP and phosphite both retain the ability to increase the stability of E_c•S (and E•$S^{\ddagger}$) relative to E_oS (red arrows in **Figure 17**)

As a result of the equivalent effects of the substitutions on the values of the values of k_{cat}/K_M for OMP and EO, the IBEs for the 5'-phosphate group for the single and most of the double mutants are unchanged from that measured for the wild type enzyme; the same observation was made for Ala substitutions for the residues in the hydrophobic cluster (19). [For the T159V/V182A and T159V/V182A/Y206F mutants, the values of the IBEs are modestly decreased, likely the result of the same small structural changes that may affect the values of k_{cat}.] We conclude that although the 5'-phosphate group provides the same IBE, E_c•S and E•$S^{‡}$ are less stabilized for the mutants than the wild type enzyme because the interactions that stabilize the closed conformation of the active site loop (Val 182) and stabilize the domain interface (Thr 159, Arg 160, and Tyr 206) have been disrupted.

Kinetic Constants for Mutants of Residues that Generate the IBE of the 5'-Phosphate Group. Using both OMP and EO, we measured the values of kinetic constants for Q185A, R203A, and Q185A/R203A, the two single mutants and one double mutant with substitutions for the two residues with side chains that directly contact the 5'-phosphate group of the substrate (**Tables 4 and 5**). With OMP, the values of k_{cat} and K_M and, therefore, k_{cat}/K_M could be measured for Q185A and R203A; however, only the value k_{cat}/K_M could be measured for Q185A/R203A.

The values of k_{cat} for the Q185A and R203A mutants are "unchanged" from that measured for wild type enzyme (**Table 4**), so the energy difference between the E_c•S and E_c•$S^{‡}$ complexes is not affected by the decreased ability of the "remote" 5'-phosphate group to activate the enzyme. Therefore, these substitutions alter the

energies of both E_c•S and E_c•$S^{‡}$ relative to E_o + S, i.e., *the 5′-phosphate group does not preferentially stabilize the transition state complex, E_c•$S^{‡}$, but also stabilizes the Michaelis complex, E_c•S, because both share the E_c conformation*.

The values of k_{cat}/K_M for OMP are decreased but the values for EO are "unchanged" from that measured for wild type. The decreased value for OMP is explained by the weaker affinity of enzyme for the 5′-phosphate group, i.e., a decrease in the value of K_s (green arrow in **Figure 17**). That the values for EO are unchanged provides evidence that Gln 185 and Arg 203 are not involved in determining the relative energies of E_o and E_c.[4] Thus, we hypothesize that the IBE enables the conformational reorganization of the active site loop that allows formation of the interdomain contacts involving Thr 159, Arg 160, and Tyr 206 as well as assembly of the hydrophobic cluster involving Val 182 in the active site loop, with these determining the stability of E_c relative to E_o.

The value of the IBE for the wild type enzyme is 11.4 kcal/mol (**Table 5**). For the Q185A and R203A mutants, the values are decreased by 2.1 and 4.0 kcal/mol, respectively; for Q185A/R203A, the value is decreased by 6.7 kcal/mol (approximately the sum of the those for the single mutants). But, despite the complete absence of the direct contacts of the 5′-phosphate group with side chains, the Q185A/R203A mutant is activated 2,900-fold by the 5′-phosphate group (the IBE is 4.7 kcal/mol). The 5′-phosphate group retains its interactions with the backbone amide groups of Gly 202 and Arg 203 as well as several ordered water molecules, including one that also hydrogen-bonds with Asp 20 in the "fixed" domain (**Figure 20A**).

The values of $(k_{cat}/K_M)/K_D$ for activation of the decarboxylation of EO by phosphite were also determined for the single mutants. The values are decreased significantly, as expected because the values of K_s for phosphite binding should be decreased by the absence of these direct interactions.

That the side chain contacts with the 5'-phosphate group are not entirely responsible for generating the IBE of the 5'-phosphate group contrasts with results we reported for similar mutagenesis-based experiments with the OMPDC from *Saccharomyces cerevisiae* (ScOMPDC) (35). In ScOMPDC, the 5'-phosphate group interacts with the side chains of three residues, Gln 215 and Arg 235, structural homologues of Gln 185 and Arg 203 in MtOMPDC, as well as Tyr 217 in the longer (19-residue) active site loop (**Figure 20B**). In ScOMPDC, the 5'-phosphate group is hydrogen-bonded to the backbone amide groups of Gly 234 and Arg 235, structural homologues of Gly 202 and Arg 203 in MtOMPDC. Also, in ScOMPDC, the 5'-phosphate group is hydrogen-bonded to three (not four) water molecules, two hydrogen-bonded to the backbone of the active site loop and the third hydrogen-bonded to Asp 37, the structural homologue of Asp 20 in MtOMPDC.

Kinetic analyses for single and multiple substitutions for the residues that contact the 5'-phosphate group (Q215A, Y217F, and R203A) revealed progressive decreases in the values of the IBE, with the values of k_{cat}/K_M for OMP only 12-fold greater than that for EO for the Q215A/Y217F/R235A triple mutant. An explanation for why the direct contacts with the side chains apparently are "fully" responsible for generating the IBE in ScOMPDC but not in MtOMPDC is not obvious. However, the active site loops differ in

both length and conformation; also, structures are not available for the mutants of ScOMPDC, so unlike MtOMPDC, the effects of the substitutions on the structure of the protein and active site cannot be assessed.

Kinetic Constants for Double Mutants of Residues that Stabilize E_c and Generate the IBE of the 5'-Phosphate Group. To obtain evidence that the "remote" interactions that stabilize E_c•S and E_c•$S^{\ddagger}$ relative to E_o•S and those that generate the IBE of the 5'-phosphate group by direct contacts are independent, we constructed and characterized the R203A/T159V, R203A/R160A, and R203A/V182A double mutants. We chose the R203A instead of the Q185A substitution because Gln 185 is hydrogen-bonded not only to the 5'-phosphate group but also to the OH group of Ser 127 in the "mobile" domain, i.e., Gln 185 could contribute to both stabilization of E_c (and, therefore, the value of K_c') as well as generation of the IBE; in fact, the Q185A substitution does cause a six-fold decrease in the value of k_{cat}/K_M for EO.[4] The kinetic constants for these mutants also are included in **Tables 4 and 5**.

The values of k_{cat} and K_M for OMP could not be measured for these mutants: because all of the single mutants effect the stabilities of E_c•S relative to E_o + S, the values of the K_M were expected to be and are, in fact, too large to be measured (**Table 4**). The values of k_{cat}/K_M for EO for the double mutants are essentially identical to those measured for the T159V, R160A, and V182A single mutants that alter the stability of E_c •S relative to E_o, i.e., the value of k_{cat}/K_M for EO for the R203A mutant is "unchanged" (only two-fold decreased) from that measured for the wild type enzyme (**Table 5**). Thus, the effects of the substitutions are additive with respect to

stabilization of E_c in both E_c•S and E_c•$S^{‡}$. Also, the values for the IBE for the 5′-phosphate group are essentially the same as that measured for the R203A mutant, i.e., the effects of the substitutions on the IBE also are additive (**Table 5**).

Independent Structural Strategies to Enable and Accomplish the Stabilization of E_c. Taken together, our structural and kinetic observations provide persuasive evidence that the model presented in **Scheme 2**/**Figure 17** provides a structure-based understanding of how the IBE of the 5′-phosphate group is both generated and used to activate the enzyme. The direct interactions of the 5′-phosphate group with the side chains of Gln 185 in the active site loop and Arg 203 (as well as the backbone amide groups of Gly 202 and Arg 203 and water molecules that interact with the active site loop; green arrows in **Figure 17**) organize the active site loop (red arrows) so that the hydrophobic cluster at the base of the active site loop and the "remote" interdomain interactions can be formed to stabilize E_c•S and E_c•$S^{‡}$ relative to E_o•S (blue arrows).

This "cascade" of structurally expanding interactions, triggered by the binding of OMP to E_o and likely orchestrated by changes in the structure of the active site loop directed by the "remote" 5′-phosphate group, results in the remarkable 2.4×10^8-fold rate enhancement as the stabilizing/destabilizing interactions are enforced in E_c•S and E_c•$S^{‡}$. Further studies of interactions of the 5′-phosphate group with its binding site, including the importance of the water-mediated contacts with the active site loop, are in progress.

Conclusions. The importance of the IBEs of "remote" substituents in achieving the transition state stabilization that are responsible for the impressive rate enhancements

of enzyme-catalyzed reactions has long been recognized (23). However, a unifying structure-based description and understanding for *how* the IBE of remote substituents is generated and used to effect transition state stabilization has been elusive. Like triose phosphate isomerase (22), MtOMPDC provides a useful system to explore and define how the IBE of the "remote" phosphate group of the substrate is generated and used to contribute to its impressive 10^{17} rate acceleration: the reaction is unimolecular and requires no cofactors, the wild type enzyme and mutants can be structurally characterized, a sensitive direct spectrophotometric assay is available, and substrate analogs (EO and FEO) can be used to "isolate" the role of the activating 5'-phosphate group. Prior structural characterization established that substrate binding in MtOMPDC is accompanied by a conformational change in which a catalytically active, closed conformation (E_c) that is unfavorable in the absence of the substrate predominates when substrate is bound. We now have obtained persuasive kinetic and structural evidence that the IBE of the 5'-phosphate group provides the energy for a local reorganization of the structure of its binding site in the E_o•S complex (in particular, the ordering of the active site loop that includes Gln 185) that then enables the formation of additional "remote" interactions at the interface between the "mobile" and "fixed" domains that provide the energy to stabilize E_c in both the E_c•S and E_c•$S^{\ddagger}$ complexes. We expect that this strategy will be applicable to understanding the structural basis for the activation of other enzymes that undergo an essential conformational change from an open, inactive conformation (E_o) to a closed, active conformation (E_c) as the result of substrate binding.

Our studies using both the intact OMP substrate and EO/FEO that lack the 5'-phosphate group (19, 25, 26, 32) provide a kinetics/structure-based approach for understanding the role of the substrate and the IBE of its pieces in enabling the conformational changes that are required for efficient enzymatic catalysis. The dynamic characteristics of conformational changes that accompany catalysis have been studied in other enzymes, including adenylate kinase (36, 37) and dihydrofolate reductase (38-40). The principles that emerge from the kinetics-based approaches described in this manuscript and previous studies of triose phosphate isomerase (22) as well as dynamics-based approaches that employ time-resolved NMR spectroscopy, isotope exchange, and simulations together 1) provide both a description and an understanding of Nature's structural strategies for achieving large rate accelerations, and 2) inform the (re)design of enzymic catalysts for novel reactions by revealing the requirements for substrate binding to E_o, utilization of energy provided by binding interactions with "remote" portions of the substrate to, in part, effect a necessary conformational transition from E_o•S to E_c•S that allows catalysis via E_c•$S^{\ddagger}$, and, finally, a favorable conformational relaxation to E_o to permit product dissociation.

ACKNOWLEDGEMENTS

We are grateful to the staff of NSLS beamlines X4A and X29A for their help with collecting diffraction data.

REFERENCES

(1) Toth, K., Amyes, T. L., Wood, B. M., Chan, K. K., Gerlt, J. A., and Richard, J. P. (2009) An examination of the relationship between active site loop size and thermodynamic activation parameters for orotidine 5'-monophosphate decarboxylase from mesophilic and thermophilic organisms. *Biochemistry 48*, 8006-13.

(2) Radzicka, A., and Wolfenden, R. (1995) A proficient enzyme. *Science 267*, 90-3.

(3) Wolfenden, R. (2011) Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annu Rev Biochem 80*, 645-67.

(4) Miller, B. G., Snider, M. J., Short, S. A., and Wolfenden, R. (2000) Contribution of enzyme-phosphoribosyl contacts to catalysis by orotidine 5'-phosphate decarboxylase. *Biochemistry 39*, 8113-8.

(5) Miller, B. G., Snider, M. J., Wolfenden, R., and Short, S. A. (2001) Dissecting a charged network at the active site of orotidine-5'-phosphate decarboxylase. *J Biol Chem 276*, 15174-6.

(6) Van Vleet, J. L., Reinhardt, L. A., Miller, B. G., Sievers, A., and Cleland, W. W. (2008) Carbon isotope effect study on orotidine 5'-monophosphate decarboxylase: support for an anionic intermediate. *Biochemistry 47*, 798-803.

(7) Wepukhulu, W. O., Smiley, V. L., Vemulapalli, B., Smiley, J. A., Phillips, L. M., and Lee, J. K. (2008) A substantial oxygen isotope effect at O2 in the OMP decarboxylase reaction: mechanistic implications. *Org Biomol Chem 6*, 4533-41.

(8) Sievers, A., and Wolfenden, R. (2002) Equilibrium of formation of the 6-carbanion of UMP, a potential intermediate in the action of OMP decarboxylase. *J Am Chem Soc 124*, 13986-7.

(9) Toth, K., Amyes, T. L., Wood, B. M., Chan, K., Gerlt, J. A., and Richard, J. P. (2007) Product deuterium isotope effect for orotidine 5'-monophosphate decarboxylase: evidence for the existence of a short-lived carbanion intermediate. *J Am Chem Soc 129*, 12946-7.

(10) Amyes, T. L., Wood, B. M., Chan, K., Gerlt, J. A., and Richard, J. P. (2008) Formation and stability of a vinyl carbanion at the active site of orotidine 5'-monophosphate decarboxylase: pKa of the C-6 proton of enzyme-bound UMP. *J Am Chem Soc 130*, 1574-5.

(11) Levine, H. L., Brody, R. S., and Westheimer, F. H. (1980) Inhibition of orotidine-5'-phosphate decarboxylase by 1-(5'-phospho-beta-d-ribofuranosyl)barbituric acid, 6-azauridine 5'-phosphate, and uridine 5'-phosphate. *Biochemistry 19*, 4993-9.

(12) Wu, N., Mo, Y., Gao, J., and Pai, E. F. (2000) Electrostatic stress in catalysis: structure and mechanism of the enzyme orotidine monophosphate decarboxylase. *Proc Natl Acad Sci U S A 97*, 2017-22.

(13) Chan, K. K., Wood, B. M., Fedorov, A. A., Fedorov, E. V., Imker, H. J., Amyes, T. L., Richard, J. P., Almo, S. C., and Gerlt, J. A. (2009) Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: evidence for substrate destabilization. *Biochemistry 48*, 5518-31.

(14) Lewis, C. A., Jr., and Wolfenden, R. (2009) Orotic acid decarboxylation in water and nonpolar solvents: a potential role for desolvation in the action of OMP decarboxylase. *Biochemistry 48*, 8738-45.

(15) Iiams, V., Desai, B. J., Fedorov, A. A., Fedorov, E. V., Almo, S. C., and Gerlt, J. A. (2011) Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: importance of residues in the orotate binding site. *Biochemistry 50*, 8497-507.

(16) Lee, J. K., and Houk, K. N. (1997) A proficient enzyme revisited: the predicted mechanism for orotidine monophosphate decarboxylase. *Science 276*, 942-5.

(17) Houk, K. N., Lee, J. K., Tantillo, D. J., Bahmanyar, S., and Hietbrink, B. N. (2001) Crystal structures of orotidine monophosphate decarboxylase: does the structure reveal the mechanism of nature's most proficient enzyme? *Chembiochem 2*, 113-8.

(18) Handschumacher, R. E. (1960) Orotidylic acid decarboxylase: inhibition studies with azauridine 5'-phosphate. *J Biol Chem 235*, 2917-9.

(19) Wood, B. M., Amyes, T. L., Fedorov, A. A., Fedorov, E. V., Shabila, A., Almo, S. C., Richard, J. P., and Gerlt, J. A. (2010) Conformational changes in orotidine 5'-monophosphate decarboxylase: "remote" residues that stabilize the active conformation. *Biochemistry 49*, 3514-6.

(20) Harris, P., Poulsen, J. C., Jensen, K. F., and Larsen, S. (2002) Substrate binding induces domain movements in orotidine 5'-monophosphate decarboxylase. *J Mol Biol 318*, 1019-29.

(21) Sullivan, S. M., and Holyoak, T. (2008) Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proc Natl Acad Sci U S A 105*, 13829-34.

(22) Richard, J. P. (2012) A Paradigm for Enzyme-Catalyzed Proton Transfer at Carbon: Triosephosphate Isomerase. *Biochemistry 51*, 2652-2661.

(23) Jencks, W. P. (1975) Binding energy, specificity, and enzymic catalysis: the circe effect. *Adv Enzymol Relat Areas Mol Biol 43*, 219-410.

(24) Wood, B. M., Chan, K. K., Amyes, T. L., Richard, J. P., and Gerlt, J. A. (2009) Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: effect of solvent viscosity on kinetic constants. *Biochemistry 48*, 5510-7.

(25) Goryanova, B., Amyes, T. L., and Richard, J. P. (2012) Phosphite Dianion Activation of Orotidine 5'-Phosphate Decarboxylase: Substituent Effects on Reactions of the Substrate Pieces. *Biochemistry*, in preparation.

(26) Amyes, T. L., Richard, J. P., and Tait, J. J. (2005) Activation of orotidine 5'-monophosphate decarboxylase by phosphite dianion: the whole substrate is the sum of two parts. *J Am Chem Soc 127*, 15708-9.

(27) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology* (Carter, C. W. J., Sweet,

R. M., Abelson, J. N., and Simon, M. I., Eds.) pp 307-326, Academic Pres, New York.

(28) Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr 64*, 125-32.

(29) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr 60*, 2126-32.

(30) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr 66*, 213-21.

(31) Lamzin, V. S., and Wilson, K. S. (1997) Automated refinement for protein crystallography. *Methods Enzymol 277*, 269-305.

(32) Barnett, S. A., Amyes, T. L., Wood, B. M., Gerlt, J. A., and Richard, J. P. (2008) Dissecting the total transition state stabilization provided by amino acid side chains at orotidine 5'-monophosphate decarboxylase: a two-part substrate approach. *Biochemistry 47*, 7785-7.

(33) Toth, K., Amyes, T. L., Wood, B. M., Chan, K., Gerlt, J. A., and Richard, J. P. (2010) Product deuterium isotope effects for orotidine 5'-monophosphate decarboxylase: effect of changing substrate and enzyme structure on the

partitioning of the vinyl carbanion reaction intermediate. *J Am Chem Soc 132*, 7018-24.

(34) Tsang, W. Y., Wood, B. M., Wong, F. M., Wu, W., Gerlt, J. A., Amyes, T. L., and Richard, J. P. (2012) Proton Transfer from C-6 of Uridine 5'-Monophosphate Catalyzed by Orotidine 5'-Monophosphate Decarboxylase: Formation and Stability of a Vinyl Carbanion Intermediate and the Effect of a 5-Fluoro Substituent. *J Am Chem Soc 134*, in press.

(35) Amyes, T. L., Ming, S. A., Goldman, L. M., Wood, B. M., Desai, B. J., Gerlt, J. A., and Richard, J. P. (2012) Orotidine 5'-Monophosphate Decarboxylase: Transition State Stabilization from Remote Protein-Phosphodianion Interactions. *Biochemistry 51*, 4630-4632.

(36) Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M., and Kern, D. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature 450*, 913-6.

(37) Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., Karplus, M., Hubner, C. G., and Kern, D. (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature 450*, 838-44.

(38) Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science 313*, 1638-42.

(39) Boehr, D. D., Dyson, H. J., and Wright, P. E. (2008) Conformational relaxation following hydride transfer plays a limiting role in dihydrofolate reductase catalysis. *Biochemistry 47*, 9227-33.

(40) Bhabha, G., Lee, J., Ekiert, D. C., Gam, J., Wilson, I. A., Dyson, H. J., Benkovic, S. J., and Wright, P. E. (2011) A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis. *Science 332*, 234-8.

(41) DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, San Carlos, CA.

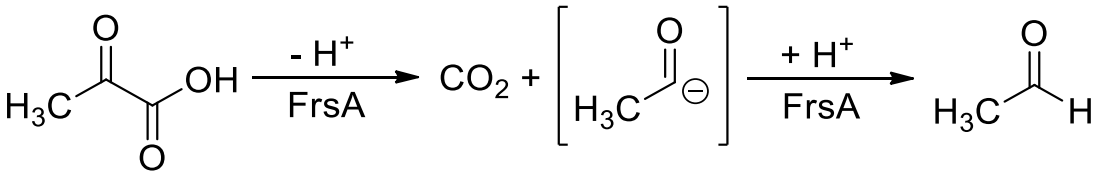
CHAPTER 4: COMPUTATIONAL, STRUCTURAL AND KINETIC EVIDENCE THAT *VIBRIO VULNIFICUS* FRSA IS NOT A COFACTOR-INDEPENDENT PYRUVATE DECARBOXYLASE

ABSTRACT: The fermentation-respiration switch (FrsA) protein in *Vibrio vulnificus* was recently reported to catalyze the cofactor-independent decarboxylation of pyruvate. We now report QM/MM calculations that examine the energetics of C-C bond cleavage for a pyruvate molecule bound within the putative active site of FrsA. These calculations suggest that the barrier to C-C bond cleavage in the bound substrate is 28 kcal/mol, which is similar to that estimated for the uncatalyzed decarboxylation of pyruvate in water at 25 ^{o}C. In agreement with the theoretical predictions, no pyruvate decarboxylase activity was detected for recombinant FrsA protein that could be crystallized and structurally characterized. These results suggest that the functional annotation of FrsA as a cofactor-independent pyruvate decarboxylase is incorrect.

A recent report identified the fermentation-respiration switch (FrsA) protein in *Vibrio vulnificus* to be a cofactor-independent pyruvate decarboxylase (**Scheme 4**) (1). Indeed, FrsA was reported to exhibit a k_{cat} of approximately 1400 s^{-1} at 37 ^{o}C, which is considerably greater than the value observed for the turnover number of the thiamin-dependent pyruvate decarboxylase from *Saccharomyces cerevisiae* (2)

Reprinted (adapted) with permission of Kellett, W. F.; Brunk, E.; Desai, B. J.; Fedorov, A. A.; Almo, S. C.; Gerlt, J. A.; Rothlisberger, U.; Richards, N. G. *Biochemistry* **2013***, 52*, 1842-1844. DOI: 10.1021/bi400093y.

This remarkable finding, if correct, would imply a significant shift from the current paradigm that Nature evolved the thiamin co-factor to generate resonance-stabilized acyl carbanion equivalents in all kingdoms of life when catalyzing the oxidative and non-oxidative decarboxylation of α-ketoacids.



Scheme 4. Cofactor-independent decarboxylation of pyruvate showing the putative acyl anion intermediate.

The X-ray crystal structure of unliganded FrsA,[1] which revealed a putative active site containing similar residues to those present in orotidine 5'-monophosphate decarboxylase (OMPDC), (3-5) provided a chemical rationale for the unexpected functional assignment of FrsA. Thus, it was argued that the catalytic power of the enzyme derived from electrostatic repulsion between pyruvate and the negatively charged side chain of Asp-203 in FrsA (1). We were intrigued by these conclusions for two reasons. First, the three-dimensional fold of FrsA places it within the α,β-hydrolase superfamily of enzymes, (6) which are known to catalyze a diverse array of reactions (7) including the decarboxylation of β-ketoacids to yield methylketones (8). OMPDC has a different fold, however, and this lack of structural similarity precludes any direct evolutionary relationship between the two enzymes (5). Second, considerable evidence

exists to suggest that stabilization of the carbanion intermediate formed in the OMPDC-catalyzed reaction is enabled by binding energy obtained from the extensive set of interactions of the protein and the sugar-phosphate moiety of the OMP substrate (9,10). On the other hand, the energy released by the interaction of the methyl substituent with FrsA and the small number of hydrogen bonds to the carbonyl group seems insufficient for stabilization of any acyl anion intermediate formed during FrsA-catalyzed decarboxylation. We therefore used advanced computational methods to evaluate the energetics of the proposed mechanism for FrsA-catalyzed conversion of pyruvate to acetaldehyde, and also prepared and assayed recombinant *Vibrio vulnificus* FrsA to determine whether the reported activity could be reproduced.

The model of the FrsA/pyruvate complex used in our computational studies was based upon the "open" monomer in the crystal structure of the free enzyme (3MVE). After adding hydrogen atoms, the protein was placed in a box of TIP3P water molecules (11) containing two chloride ions to yield a neutral system. The resulting structure was energy minimized and equilibrated by molecular dynamics (MD) simulation. Parameters for pyruvate were obtained from the generalized AMBER force field (12-14) and the substrate was docked into the putative enzyme active site using GLIDE (15). Energy minimization and MD equilibration of several model complexes with pyruvate in different orientations within the putative active site all gave the same final position for the substrate (**Figure 24**).

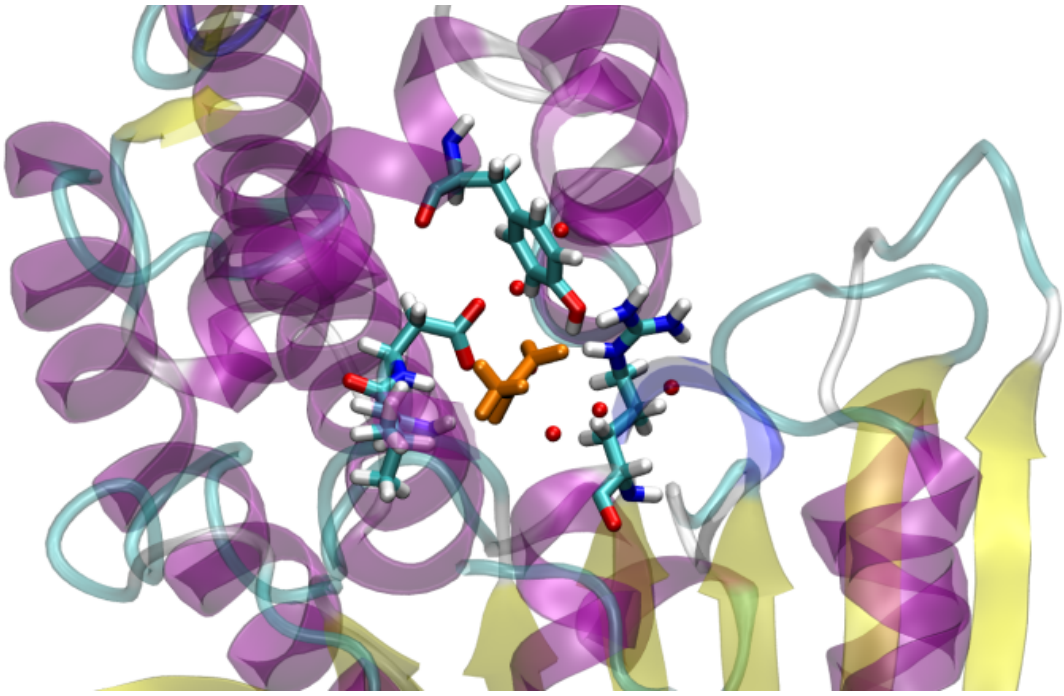


Figure 24. Equilibrated model of pyruvate (orange) docked into the putative active site of FrsA. Pyruvate is hydrogen bonding to the backbone NH of Leu-202 and the side chains of Tyr-316 and Arg -272. Color scheme: C – cyan; H – white; N – blue; O – red. Active site water oxygen atoms are rendered as red spheres.

The final equilibrated structure of the pyruvate/FrsA complex resembled the one proposed previously (1), with pyruvate forming hydrogen bonds to the side chains of the backbone NH of Leu-202, and the side chains of Arg-272 and Tyr-316. In addition, three active site water molecules associated strongly with bound pyruvate throughout these MD simulations. This solvated model of the pyruvate/FrsA complex proved to be stable in an unconstrained NPT MD simulation over a period of 20 ns, and so was used in a series of QM/MM simulations of the C-C bond cleavage reaction employing an extension of the Car-Parinello MD (CPMD) methodology (16). The QM region consisted of pyruvate, the Tyr-316 side chain up to the C_β atom, and three active site waters. These atoms were described by the BLYP functional (17,18) and norm-conserving

Martins-Trouiller pseudopotentials (19) with dispersion-corrected atom-centered dispersion potentials (20-22). The remaining atoms, comprising the rest of the protein and explicit water molecules, were described by the classical AMBER99 force field (13,14). The side chains of the hypothetical "catalytic residues, Asp-203 and Arg-272, were not included in the QM region because their putative electrostatic contributions to catalysis could be adequately represented using an MM description. In the CPMD calculations, the C1-C2 bond distance in pyruvate was chosen as the reaction coordinate; hence, constraints were employed at distances spanning 1.55 to 4.24 Å (in increments of 25 pm). The QM/MM system was equilibrated for 2 ps at constant pressure and temperature before performing constrained MD simulations for thermodynamic integration (23,24) in the NPT ensemble. Each system was sampled for 1 ps and the free energy profile was computed by integrating the constraint forces over the respective distances (**Figure 25**). These simulations gave an estimated free energy barrier of 28.1 ± 0.2 kcal/mol for the conversion of FrsA-bound pyruvate into acetaldehyde and CO_2, corresponding to a first-order rate constant of 1.1×10^{-9} s^{-1} at 25 ^{o}C, assuming transition state theory and the absence of re-crossing (25). This value is very similar to the experimental estimate of the first order rate constant for the uncatalyzed decarboxylation of pyruvate, which has an upper limit of approximately 10^{-9} s^{-1} at this temperature and pH 7 (26). The calculated value should be considered as a lower bound given that BLYP is known to underestimate activation barriers, especially those for proton transfer steps (29). For example, "benchmark" studies give an estimate of 23.0 ± 3.1 kcal/mol for the uncatalyzed reaction in water. Certainly both

computed barriers are inconsistent with the reported[1] turnover number of 1400 s^{-1} for FrsA-catalyzed decarboxylation.

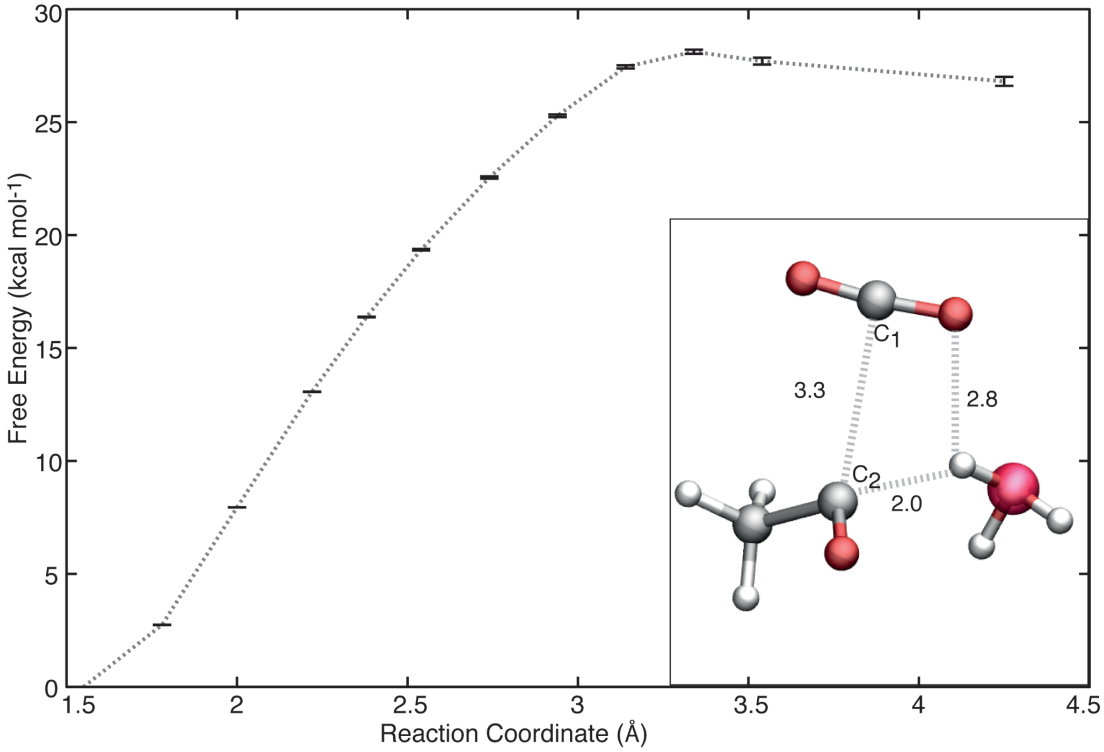


Figure 25. Free energy profile (kcal/mol) for cleavage of the C1-C2 bond in FrsA-bound pyruvate, as computed by thermodynamic integration. Error bars show the statistical sum of errors associated with the calculated free energy. The inset shows the active site configuration at the transition state (C-C = 3.3 Å) and the transfer of a proton from a hydronium ion to the acyl anion.

Furthermore, the notion that the FrsA protein environment does not catalyze C-C bond cleavage is consistent with the fact that the QM/MM simulations suggest that a nearby tyrosine residue (Tyr-316) protonates pyruvate during the reaction mechanism (**Figure 26** and **Appendix C**). This proton is subsequently transferred to a nearby water molecule at the transition state (C-C bond length of 3.1 Å). The resulting hydronium ion stabilizes the developing anionic charge on the central carbon atom when the C-C bond length in the substrate is elongated to 3.3 Å, and protonates the

acyl anion, thereby giving acetaldehyde (**Figure 26**). Both the initial deprotonation of tyrosine by substrate and the subsequent proton transfer to water are counter-intuitive on the basis of standard pK_a values, and calculations[28] do not support any large pK_a shifts for either Tyr-316 or bound pyruvic acid (**Appendix C**). These computational findings again point to the fact that FrsA is unlikely to be a co-factor independent decarboxylase.

To evaluate the conclusions of the QM/MM and pK_a shift calculations, the gene encoding *Vibrio vulnificus* FrsA was expressed in *Escherichia coli* and purified *via* two different procedures. In agreement with the computational results, neither preparation exhibited detectable pyruvate decarboxylase activity (acetaldehyde production or CO_2 evolution) in a coupled-enzyme spectrophotometric assay (alcohol dehydrogenase), by ^1H NMR spectroscopy, or by membrane-inlet mass spectrometry (**Appendix C**) (29). The purified protein was, however, crystallized and its structure determined at 1.95 Å resolution; our structure was essentially identical to that reported earlier except an unknown ligand, modeled as hexanoate, was located in the putative active site (**Appendix C**, **Figure 38**). Given that α,β-hydrolase superfamily members are functionally diverse, (8) it is difficult to predict *in vitro* activity solely from sequence homology. However, these computational, structural, and experimental results do not support the claim that FrsA is a cofactor-independent pyruvate decarboxylase.

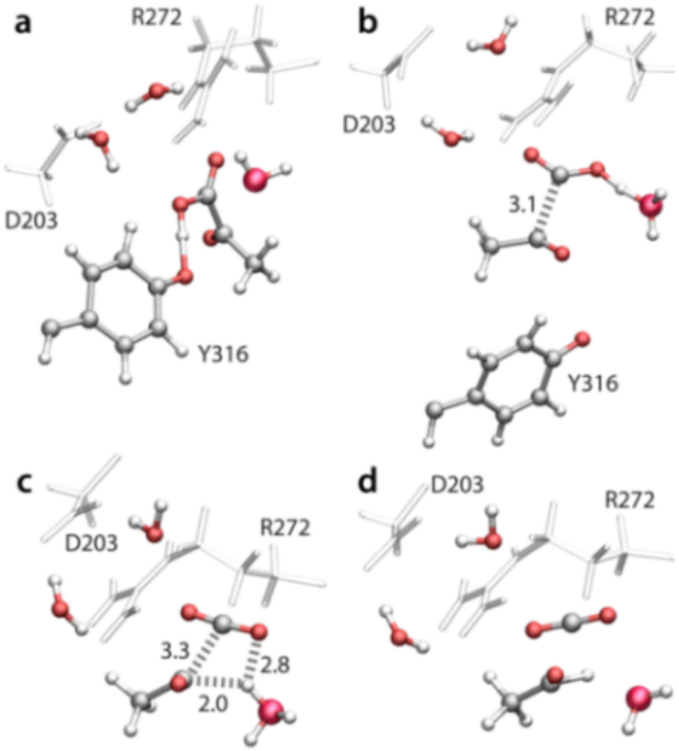


Figure 26. Molecular events observed during the cleavage of the C1-C2 bond in FrsA-bound pyruvate. (a) Protonation of the substrate carboxylate by the side chain of Tyr-316. (b) Deprotonation of the carboxylic acid via a nearby water molecule. (c) Formation of a hydronium ion during decarboxylation, which then acts to quench the developing anionic charge on C2 thereby yielding acetaldehyde (d). Atoms in the QM and MM regions are rendered as "ball-and-stick" and licorice representations, respectively. Proton transfers are shown using "dynamic bonds". Color scheme for the QM atoms: C, grey; H, white; O, red.

Funding Sources

These studies were funded by the Swiss National Science Foundation (U.R.), the National Institutes of Health (DK061666 to N.G.J.R., and GM065155 to J.A.G. and S. C. A.), and a Center for Synchrotron Biosciences grant (P30-EB-009998) from the National Institute of Biomedical Imaging and Bioengineering. Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department

of Energy (contract DE-AC02-98CH1088). The University of Florida also provided an Alumni Fellowship (W.F.K.).

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENTS

We thank Richard Wolfenden for discussions concerning the rate of uncatalyzed pyruvate decarboxylation in water, and Sun-Shin Cha who provided the pQE-FrsA plasmid and useful information about FrsA protein production and assay methods. We are also grateful to the staff of NSLS beamline X29A for their help with diffraction data collection.

REFERENCES

(1) Lee, K.-J., Jeong, C.-S., An, K. J., Lee, H.-J., park, S.-J., Seok, Y.-J., Kim, P., Lee, J.-H., Lee, K.-H., and Cha, S.-S. (2011) *Nat. Chem. Biol. 7*, 434-436.

(2) Sergienko, E. A., and Jordan, F. (2001) *Biochemistry 40*, 7369-7381.

(3) Miller, B. G., and Wolfenden, R. (2002) *Annu. Rev. Biochem. 71*, 847-885.

(4) Chan, K. K., Wood, B. M., Fedorov, A. A., Fedorov, E. V., Imker, H. J., Amyes, T. L., Richard, J. P., Almo, S. C., and Gerlt, J. A. (2009) *Biochemistry 48*, 5518-5531.

(5) Wu, N., Mo, Y. R., Gao, J. L., and Pai, E. F. (2000) *Proc. Natl. Acad. Sci., USA 97*, 2017-2022.

(6) Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G., and Goldman, A. (1992) *Prot. Eng. 5*, 197-211.

(7) Bugg, T. D. H. (2004) *Bioorg. Chem. 32*, 367-375.

(8) Auldridge, M. E., Guo, Y., Austin, M. B., Ramsey, J., Fridman, E., Pichersky, E., and Noel, J. P. (2012) *Plant Cell 24*, 1596-1607.

(9) Goryanova, B., Amyes, T. L., Gerlt, J. A., and Richard, J. P. (2011) *J. Am. Chem. Soc. 133*, 6545-6548.

(10) Amyes, T. L., Richard, J. P., and Tait, J. J. (2005) *J. Am. Chem. Soc. 127*, 15708-15709.

(11) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys. 79*, 926-935.

(12) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) *Proteins 61*, 712-725.

(13) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) *J. Comput.Chem. 25*, 1157-1174.

(14) Wang, J., Cieplak, P., and Kollman, P. A., (2000) *J. Comput. Chem. 21*, 1049-1074.

(15) Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shaw, D. E., Shelley, M., Perry, J. K., Francis, P., and Shenkin, P. S. (2004) *J. Med. Chem. 47*, 1739-1749.

(16) Laio, A., VandeVondele, J., and Rothlisberger, U. (2002) *J. Chem. Phys.* 116, 6941–6948.

(17) Lee, C., Yang, W., and Parr, R. G. (1988) *Phys. Rev. B. 37*, 785–789.

(18) Becke, A. D. (1988) *Phys. Rev. A. 38*, 3098–3100.

(19) Trouiller, N., and Martins, J. L. (1991) *Phys. Rev. B. 43*, 8861–8869.

(20) Arey, J. S., Aeberhard, P. C., Lin, I.-C.,and Rothlisberger, U. (2009) *J. Phys. Chem. B 113*, 4726–4732.

(21) Lin, I.-C., Coutinho-Neto, M. D., Felsenheimer, C., von Lilienfeld, O. A., Tavernelli, I., and Rothlisberger, U. (2007) *Phys. Rev. B 75*, 205131.

(22) von Lilienfeld, O. A., Tavernelli, I., Rothlisberger, U., and Sebastiani, D. (2007) *Phys. Rev. Lett. 93*, 153004.

(23) Sprik, M., and Ciccotti, G. (1998) *J. Chem. Phys. 109*, 1739–1749.

(24) Kirkwood, J. G. (1935) *J. Chem. Phys. 3*, 300-313.

(25) Himo, F. (2006) *Theor. Chem. Acc. 116*, 232-240

CHAPTER 5: GENERAL CONCLUSIONS

This work has investigated several aspects of OMPDC that are crucial for achieving its remarkable catalytic efficiency. Chapter 2 investigates the mode by which the vinyl carbanion is stabilized. Our studies shows that interaction with O4 of the substrate is not the major means of stabilization of the anion (via the carbene resonance structure). Several interactions with the active site, especially interaction with Lys72, play a role in transition-state stabilization. Moreover, computational calculations show that the negative charge produced in the transition-state is more diffusely distributed over the pyrimidine ring, contrary to earlier understanding that the negative charge resides at a particular locus, like O4 or C6 of the pyrimidine ring. Therefore multiple interactions within the active site are responsible for stabilization of the vinyl carbanion intermediate. Our studies also suggest that transition-state stabilization cannot solely account for the 10^{17} fold rate enhancement, and thus indirectly finding support for ground-state destabilization as a means of reducing the activation energy barrier for OMP decarboxylation.

Our studies on residues remote from the active site in Chapter 3 (T159, V182 and Y206), show that these residues can stabilize the catalytically competent closed conformation of the enzyme (E_c). Mutation of these residues can increase the K_M by shifting the equilibrium between E_o and E_c towards E_o. This discovery, that K_M is a

complex that reflects both binding equilibrium and equilibrium between conformational states, enriches the current understanding in interpreting kinetic data from enzymes.

Investigation of phosphate-binding interactions indicates that a majority (~7 kcal out of 11.4 kcal), but not all, of the intrinsic phosphate-binding energy is furnished by interaction of R203 and Q185 with the 5'-phosphate group. Moreover, our studies provide a structural explanation for how intrinsic binding energy can enable a conformational change to aid catalysis. Prior to this work, intrinsic phosphate binding energy was proposed to provide free energy for catalysis by providing additional interactions that manifest only in the transition-state. There was no structural evidence for this. Our studies show that intrinsic binding energy is generated by binding interactions with residues in the active site of OMPDC in both the ground-state and transition-state. These interactions pay for entropic costs of desolvation, enthalpic penalty due to GSD and enabling the formation of the catalytically competent Michaelis complex or E_c, possibly by changing the conformational landscape of the enzyme. This unique role of the remote 5'-phosphate group is important for enforcing strategies required for efficient cofactor-less decarboxylation of OMP.

In Chapter 4, I show that FrsA cannot catalyze cofactor-less decarboxylation of pyruvate and therefore is misannotated as a pyruvate decarboxylase. Pyruvate does not have a phosphate handle that can provide a large intrinsic binding energy to perform the difficult catalytic reaction of pyruvate decarboxylation. Moreover, the active site of

FrsA does not provide any stabilization of the carbanion formed after the decarboxylation.

Taken together, this thesis adds a deeper, structure-based understanding of the enigmatic efficiency of OMPDC. Principles gained through this work, on strategies for stabilization of anionic intermediates and use of remote-substituent binding energy to drive catalysis, can be used to guide enzyme engineering/re-engineering efforts.

APPENDIX A: SUPPORTING INFORMATION FOR CHAPTER 2

INVESTIGATING THE ROLE OF A BACKBONE TO SUBSTRATE HYDROGEN BOND IN OMP DECARBOXYLASE USING A SITE-SPECIFIC AMIDE TO ESTER SUBSTITUTION

Bijoy J. Desai[a,c], Yuki Goto[d], Alessandro Cembran[e], Alexander A. Fedorov[f], Stephen C. Almo[f], Jiali Gao[e], Hiroaki Suga[d] and John A. Gerlt[a,b,c]

[a]Department of Biochemistry and [b]Department of Chemistry, 600 South Mathews Avenue, and [c]Institute for Genomic Biology, 1206 West Gregory Drive, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States
[d]Department of Chemistry, Graduate School of Science, University of Tokyo, Tokyo 113-0033, Japan
[e]Department of Chemistry, University of Minnesota, 207 Pleasant Street SE, Minneapolis, MN 55455, United States
[f]Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY, United States

TABLE OF CONTENT

Supplemental Methods

All reagents were purchased from Sigma-Aldrich unless stated otherwise. [^{35}S]-Methionine was purchased from Perkin-Elmer. All solutions were made in ultra-pure deionized water.

Plasmids Preparation. The MtOMPDC gene optimized for *Escherichia coli* codon usage and formation of minimal secondary structure in mRNA transcripts was synthesized by Genescript and cloned into the plasmid supplied with NEB PURExpress® *in vitro* translation kit between NdeI and BamHI restriction sites (**Table 12**). Site-directed mutagenesis using overlap extension was used to first generate H128N mutation by previously established procedures (1-3). The MtOMPDC H128N plasmid was used to perform a second round of site-directed mutagenesis to generate gene containing MtOMPDC S127Amber/H128N mutation (**Table 14**). For heterologous expression of MtOMPDC H128N from *E. coli*, the gene was cloned into expression vector PET15b between NdeI and BamHI restriction sites.

Purification of MtOMPDC H128N and S127Q from *Escherichia coli*. MtOMPDC H128N was overexpressed in *E. coli* BL21 (DE3) and MtOMPDC S127Q was overexpressed in *E. coli* BW25113 OMPDC (*pyrF)* K/O strain, using previously published procedure, with slight modifications (1-3). Overnight starter cultures of the *E.coli* cells transformed with PET15b plasmid containing the MtOMPDC gene were inoculated into 6

X 2 L Luria-Bertani (LB) medium containing appropriate antibiotics. The cells were incubated for 12-14 hrs with aeration. The cells were harvested by centrifugation at 4000 g for 10 min. The harvested cells were re-suspended in equilibrating buffer (20 mM Tris HCl, pH 7.9) and lysed by sonication. The cell lysate was separated by centrifugation at 31,000 g. The supernatant lysate was charged onto a pre-equilibrated column containing Ni-Sepharose resin (GE Life sciences). The column was washed with wash buffer (20 mM Tris, pH 7.9, 500 mM NaCl, 50 mM imidazole) and the bound His-tagged protein was eluted using elution buffer (20 mM Tris, pH 7.9;,500 mM NaCl, 1 M imidazole) by gradient elution. The fractions containing OMPDC (determined by SDS-PAGE) were pooled and dialyzed against 1 X PBS. The His-tag was cleaved using thrombin (GE Life sciences) at 25° C for 12 hrs. The completion of the cleavage reaction was confirmed using SDS-PAGE. The cleaved protein was separated using FPLC on a column containing phenyl Sepharose resin (GE Life sciences) pre-equilibrated with 20 mM Tris pH 7.9, 1 M $(NH_4)_2SO_4$. Ammonium sulfate (4M) was added to the protein solution to give a final concentration of 1 M. This solution was charged onto the phenyl Sepharose column. The protein was eluted by gradient elution using 20 mM, Tris pH 7.9, as the elution buffer. The fractions containing the cleaved MtOMPDC protein were pooled and dialyzed against storage buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM DTT). The dialyzed protein was concentrated to ~17 mg/mL using Millipore Ultracell concentrators. The concentrated protein was aliquoted and flash frozen and stored at -80°C for future use.

OMP Decarboxylation Assay of Purified MtOMPDC Wt and H128N. OMP decarboxylation assays were carried out using previously published procedure (1-3), at 25°C in 10 mM MOPS pH 7.5 containing 100 mM NaCl by following the decay in absorbance at 279 nm due to formation of UMP ($\Delta\varepsilon_{279}$= 2400 $M^{-1}cm^{-1}$). OMP stock concentration was determined spectrophotometrically at 25°C in 0.1 N HCl, using molar extinction coefficient (ε_{267}= 9430 $M^{-1}cm^{-1}$). Enzyme stock concentration was determined at 25°C in 10 mM MOPS pH 7.5 containing 100 mM NaCl using the molar extinction coefficient (ε_{280}= 5960 $M^{-1}cm^{-1}$). A final enzyme concentration of 80 nM was used in the assays. k_{cat} was determined by measuring initial velocity (V_o) using saturating concentrations of OMP (where V_o is independent of [OMP]). K_M was determined by following the first order decay at low OMP concentrations (linear dependence of V_o on [OMP]). This curve was fit to a first order decay function and the first order constant was determined using non-linear regression. This constant represents V_{max}/K_M. The K_M was calculated from V_{max}/K_M using the enzyme concentration and k_{cat} that had been previously determined. The kinetic constants for MtOMPDC Wt and H128N are given in **Table 6**.

Preparation of reagents used for *in vitro* tRNA acylation using Flexizyme

Preparation of Flexizyme: Flexizyme dFx was prepared using *in vitro* run-off transcription of template DNA encoding dFx using previously published procedures (4, 5). The sequence of dFx is given in **Table 15**.

Preparation of amber suppressor tRNA: Orthogonal suppressor tRNA containing an amber anti-codon was prepared using *in vitro* run-off transcription of DNA template encoding the tRNA using previously published procedure (4). The tRNA sequence is given below in **Table 15**.

Preparation of activated Ser and ^{HO}Ser: Ser-DBE was synthesized with the same procedure as previously described (6). ^{HO}Ser-DBE was synthesized using the following procedure: L-Glyceric acid hemicalcium salt (180 mg, 1.25 mmol) was dissolved in 1 mL of 1 M HCl aq. saturated with NaCl and extracted with 4 mL of THF (5 times). The combined organic phase was dried over $MgSO_4$ and evaporated. The residue (L-glyceric acid, 122 mg, 1.15 mmol) was mixed with 3,5-dinitrobenzyl chloride (224 mg, 1.03 mmol) and triethylamine (153 µL, 1.09 mmol) in 0.3 mL of DMF, and the reaction mixture was stirred at room temperature for 16 h. After the reaction, Et_2O (9 mL) was added and the solution was washed with 1 M HCl (3 mL x 3), sat. $NaHCO_3$ (3 mL x 3) and brine (5 mL x 1), and the organic layer was dried over $MgSO_4$ and concentrated under reduced pressure. The crude residue was purified by silica gel column chromatography to give ^{HO}S-DBE (197 mg, 55%): ^{1}H NMR ($CDCl_3$, 300 MHz) δ 9.03 (t, J = 2.1 Hz, 1H), 8.59 (d, J = 2.1 Hz, 2H), 5.47 (m, 2H), 4.42 (br, 1H), 3.99 (br, 2H), 3.17 (br, 1H), 2.29 (br, 1H).

Acylation of amber suppressor tRNA using Flexizyme. Acylation of amber suppressor tRNA was achieved by incubation of 25 µM tRNA, 25 µM flexizyme and 5 mM activated L-serine (S) or L-glycerate (^{HO}S) at 0 °C in 50 mM HEPES-KOH, pH 7.5,

containing 600 mM $MgCl_2$. Before starting the reaction, the components were mixed at room temperature, except $MgCl_2$ and S-DBE or ^{HO}S-DBE. The mixture was heated at 95 °C for 3 min in order to denature the flexizyme and tRNA. The mixture was cooled to room temperature before adding $MgCl_2$ and activated S-DBE or ^{HO}S-DBE. The reactions were incubated at 0 °C for 12 hours.

After incubation, the acylated tRNA was precipitated by adding 11 volumes of a 3:8 mixture of 0.3 M sodium acetate pH 5.2, and 70% ethanol. The precipitated tRNA was washed twice using 5 volumes of 70% ethanol containing 0.1 M sodium acetate, pH 5.2, and once using 1 volume of 70% ethanol. The resulting pellet was dried at room temperature for 5 min and resuspended in RNase free deionized water immediately before use.

Quantification of *in vitro*-Synthesized Proteins. Wild type (H128N) protein synthesized by *in vitro* translation was quantified by measuring the initial rates of OMP decarboxylation at saturating substrate concentrations (V_{max}). These rates were interpolated on a standard curve of V_{max} vs. enzyme concentration generated by known quantities of H128N protein expressed in and purified form *Escherichia coli*. The amounts of S127^{HO}S, the H128N positive control (suppression with L-Ser), and negative control (S127Q) were quantified using autoradiography by comparing the intensities of the full-length, [^{35}S]-methionine labeled proteins with that of the wild type (H128N) protein (**Figure 34**).

Autoradiography of *in vitro*-Synthesized Proteins. An aliquot (2.5 µL) of the *in vitro* reaction performed in presence of [^{35}S]-methionine was mixed with equal volume of SDS loading buffer. 4.5 µL of the resulting mixture was loaded on a 4-20% polyacrylamide gradient gel (Biorad). The gel was placed in an electrophoretic cell (Biorad) containing SDS PAGE Tris-glycine running buffer and allowed to separate at constant voltage of 200 V for 30-35 min. The gel was dried using a vacuum drier (Biorad) and used to expose storage phosphorimager cassettes over-night in the dark. The cassettes were imaged using Molecular Dynamics Storm phosphorimager. The image was analyzed using ImageQuant software and the relative intensity of each protein band was determined taking wild type (H128N) as 100% (**Figure 34**). The concentration of *in vitro* synthesized wild type (H128N) determined as described previously and the relative intensities of the synthesized proteins were used to calculate the concentrations of S127^{HO}S and H128N positive control.

Alkali hydrolysis treatment of *in vitro* reaction (Chapter 2, **Figure 11*B***) was performed by adding 1 µl of 1 M $NaHCO_3/Na_2CO_3$ buffer, pH 10, to 2.5 µl of each reaction and heating at 70°C for 1 hr in a PCR thermocycler, before performing SDS-PAGE experiments.

Crystallization and Data Collection of MtOMPDC H128N. The crystals of MtOMPDC.H128N•BMP (**Table 16**) were grown by the sitting drop method at room temperature. The protein solution contained the MtOMPDC mutant (17 mg/mL) in 20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM DTT, and 40 mM 6-hydroxyuridine-5′-

phosphate (BMP); the precipitant contained 20% PEG 8000, 0.1 M Tris, pH 8.5, and 0.2 M $MgCl_2$. Crystals appeared in 7-9 days and exhibited diffraction consistent with the space group $P2_1$, with two polypeptides per asymmetric unit.

Prior to data collection, the crystals of MtOMPDC.H128N•BMP were transferred to cryoprotectant solutions composed of their mother liquids and 40% ethylene glycol and flash-cooled in a nitrogen stream. The X-ray diffraction data set (**Table 16**) was collected at the NSLS X29A beamline (Brookhaven National Laboratory) on the 315Q CCD detector. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK (7). The data collection statistics are given in **Table 16**.

Structure Determination and Model Refinement. The MtOMPDC.H128N•BMP structure was solved by molecular replacement with fully automated molecular replacement pipeline BALBES (8), using only input diffraction and sequence data. The previously determined MtOMPDC•BMP structure (PDB ID 3LTP) was used by BALBES as a search model. The partially refined structure of MtOMPDC.H128N•BMP was the output from BALBES. Subsequent several iterative cycles of refinement were performed including: model rebuilding with COOT (9), refinement with PHENIX (10), and automatic model rebuilding with ARP (11).

The quality of the final structures was verified with omit maps. The stereochemistry was checked with WHATCHECK (12) and MOLPOBITY (13). The program LSQKAB (14) was used for structural superposition. Structural figures with electron density maps were prepared using PYMOL (15). All loops are well-defined in

the MtOMPDC.H128N•BMP structure; the electron densities for both BMP ligands in the protein dimer are well-ordered; no non-glycine residue lie in disallowed regions of the Ramachandran plot. Two protein monomers within dimer are connected by a non-crystallographic two-fold axis. Final crystallographic refinement statistics are shown in **Table 16**.

Computational Methods

The wild-type OMPDC simulation system was constructed based on the crystal structure of the BMP-bound complex of MtOMPDC (PDB code: 1LOR) (16) as a template. In particular, the BMP in the "B" subunit was converted to the protonated reaction product uridine monophosphate (UMP), whereas the BMP in the "A" subunit was converted to orotidine monophosphate (OMP), which is treated by using a hybrid quantum mechanical and molecular mechanical (QM/MM) approach. The orotate group was included in the QM region, and the ribose ring C1' atom was defined as the boundary separating the QM region and the MM fragment and treated by the generalized hybrid orbital (GHO) method (17) (Chapter 2, **Figure 13*B***). All the protein's histidine residues were considered fully protonated; sodium ions were added to ensure global electroneutrality.

The protein and the ribosylphosphate group in the MM region were represented by the CHARMM 22 force field (18) along with the TIP3P model for water. The QM region was described by the semiempirical AM1 method. Previous studies have shown that the performance of AM1 for the OMP decarboxylation reaction in the gas phase is excellent and in quantitative agreement with high-level ab initio molecular orbital calculations (CBS) and with density functional theory. The Lennard-Jones parameters for atoms in the QM region (19) had been optimized for the AM1 model in QM/MM calculations. All MD simulations were performed at 298.15 K and 1 atm, controlled through the Nosé-Hoover thermostat algorithm and the Andersen piston scheme (20-22). A time-step of 1 fs was used together with the SHAKE (23) (algorithm to constrain

all the covalent bond distances involving non-QM hydrogen atoms. The TIP3P water model was used to fill a tetragonal box of 69x69X95 $Å^3$; long-range interactions were treated with particle mesh Ewald (24) (PME) with a 1 Å mesh. Lennard-Jones interactions were switched in the 10 - 11 Å range; the nonbonded lists were updated every 25 steps. Before production runs, the system was gradually heated from 20 K to 298.15 K over 40 ps and then further equilibrated for 35 ps. To study the $S127^{HO}S$ substitution, we mutated *in silico* Ser 127 to L-glycerate in the equilibrated structure, further followed by a short equilibration period. To study the effects of the $S127^{HO}S$ substitution on the decarboxylation reaction, the backbone of the S127 amide group was replaced by an oxygen atom using a structure corresponding to the Michaelis complex from the wild-type OMPDC simulations.

The umbrella sampling (25) technique was used to compute the potential of mean force (26) (PMF) along the decarboxylation reaction coordinate (Rc), defined as the bond distance between the C6 and C6' atoms that is cleaved in the decarboxylation process. To drive the sampling through the Rc range spanning from 1.4 Å to 3.0 Å, a force constant of 20 kcal mol^{-1} $Å^{-2}$ was used together with a biasing potential estimated from Reference 27 (27). Each window was subject to 30 ps of equilibration followed by 180 ps of sampling. Separate simulations, called simulation windows, were performed with the harmonic potential placed evenly with 0.2 Å interval in the 1.5 - 2.7 Å range along the reaction coordinate. One additional window was placed at 3.0 Å. A mesh of 0.02 Å was used to collect the data to obtain the distribution of reaction coordinate values and the WHAM (28) algorithm was used to reconstruct the PMF.

Trajectories were saved every 100 fs and subject to further analysis. The distances between the orotate O4 oxygen and the amide nitrogen of Ser 127, or the ester oxygen of glycerate 127 were mapped in the transition state windows and their probability densities calculated and plotted. To decompose the contribution of Ser 127 and its mutant to transition state stabilization/ground state destabilization, we computed the interaction energy between this residue and the QM region of the substrate in the reactant (R), transition state (TS), and product (P) regions. In the scheme employed, the QM energy of the substrate was first computed in the full enzyme electrostatic environment, and then it was computed again after the point charges of residue 127 were zeroed. The energy difference between these two computations is defined as the interaction energy between substrate and residue 127. The same approach was employed to calculate energetic contributions only from the side-chain of residue 127, from which the backbone interaction energy was derived by taking the difference. In all cases the MM atom selections whose charges were zeroed were globally neutral.

Mulliken Charges Analysis. The charge distribution of the substrate of the QM region was monitored along the entire reaction by Mulliken population analysis. Similar calculations have been performed using an electrostatic potential (ESP) fitting procedure both at the Hartree-Fock and MP2 level (results not shown), and the trends are consistent with the population analysis. In **Table 9** and **Table 10** the charges for the Wild-Type (WT) and for the S127Glycerate mutant (MUT) are reported at reaction

coordinates corresponding to the Reactant (R), Transition State (TS), and Product (P) states. In the active site of the wild-type OMPDC, the carbanion formal charge of the decarboxylation product is delocalized throughout the entire pyrimidine ring, with local net charges of -0.221 a.u. at C6, -0.358 a.u. on C5H, and -0.255 a.u. on C4=O4, with the remaining charge density spread out at other sites. Overall, the charge distribution analyses did not show major differences between WT and MUT. The most significant change is a shift of electron density of 0.03 a.u. from the C4=O4 carbonyl oxygen towards the C6 carbanion. Notice that the net charge on the C5H unit is not affected by the S127Glycerate replacement. Nevertheless, the change of the electrostatic environment in the mutant affects the free energy barrier (4.7 kcal/mol), and interaction energy analyses show that this change is mainly due to the loss of stabilization of the TS deriving from S127 backbone hydrogen bond with O4.

Supplemental Tables

Table 6. Kinetic parameters of OMP decarboxylation for MtOMPDC wild type and H128N, measured at 25°C and pH 7.1

	k_{cat} (sec^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}sec^{-1}$)
Wt	5.3 ± 0.2	1.6 ± 0.5	3.3×10^6
H128N	3.4 ± 0.24	2.6 ± 0.4	1.3×10^6

Table 7. FOMP decarboxylation assay of various proteins produced by *in vitro* expression, carried out at 25°C and pH 7.1

		Volume of *in vitro* reaction in assay	[FOMP] μM	V_o (M/s) X 10^{-8}
a	No DNA	10	455	0.5
b	No tRNA	10	455	0.7
c	L-glycerate suppressed	10	455	11
d	L-glycerate suppressed	10	685	10
e	L-glycerate suppressed	10	113	4.1
f	L-Ser suppressed	10	455	320
g	Wt (H128N)	5	455	1026

Table 8. Kinetic parameters of FOMP decarboxylation for MtOMPDC H128N and S127Q measured at 25°C and pH 7.1

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
H128N	340 ± 40	90 ± 10	(3.8 ± 0.6) x 10^6
S127Q	-	-	244 ± 19

Table 9. Average Mulliken population charges of the pyrimidine ring of OMP substrate at the Michaelis complex reactant state (R), at the decarboxylation transition state (TS), and at the decarboxylation product state (P) in the active site of wild-type OMPDC.

WT	R	TS	P
N1	-0.217	-0.213	-0.222
C6	0.089	-0.146	-0.221
C2	0.427	0.384	0.363
O2	-0.358	-0.385	-0.396
N3	-0.358	-0.377	-0.389
H3	0.280	0.269	0.260
C4	0.364	0.335	0.316
O4	-0.484	-0.532	-0.571
C5	-0.364	-0.456	-0.510
H5	0.177	0.170	0.152
C6'	0.340	0.478	0.451
OC1	-0.551	-0.359	-0.211
OC2	-0.548	-0.364	-0.251
C-Me	0.292	0.291	0.292

Table 10. Average Mulliken population charges for the pyrimidine group of OMP at the reactant state (R), transition state (TS), and product state (P) in the active site of $S127^{HO}S$.

MUT	R	TS	P
N1	-0.215	-0.209	-0.209
C6	0.095	-0.171	-0.257
C2	0.436	0.395	0.377
O2	-0.353	-0.381	-0.396
N3	-0.354	-0.370	-0.381
H3	0.279	0.267	0.260
C4	0.365	0.333	0.315
O4	-0.475	-0.515	-0.541
C5	-0.350	-0.456	-0.506
H5	0.174	0.164	0.148
C6'	0.354	0.473	0.452
OC1	-0.603	-0.372	-0.207
OC2	-0.549	-0.361	-0.249
C-Me	0.288	0.287	0.289

Table 11. Interaction energy (in kcal mol^{-1}) of the QM region of OMP with residue 127 in the reactant (R), transition state (TS), and product (P) regions of the reaction coordinate. The residue interaction energy is further separated in its side chain and backbone components.

	Residue	Side Chain	Backbone
WT-R	-5.1	-0.8	-4.3
WT-TS	-6.1	0.1	-6.2
WT-P	-6.8	1.6	-8.4
$S127^{HO}S$-R	5.8	-1.8	7.6
$S127^{HO}S$ - TS	7.7	-1.3	9.0
$S127^{HO}S$ -P	10.2	-0.2	10.5

Table 12. MtOMPDC optimized gene sequence

>MtOMPDC_optimized
ATGAACCGCCTGATTCTGGCTATGGACCTGATGAACCGTGATGATGCTCTGCGTGTTACCGGT GAAGTGCGTGAATACATCGACACTGTGAAAATTGGTTATCCGCTGGTACTGTCTGAAGGCATG GATATCATTGCTGAATTCCGTAAACGCTTCGGCTGCCGTATCATTGCTGACTTCAAAGTGGCA GATATCCCGGAAACCAACGAAAAAATTTGTCGTGCGACTTTCAAAGCTGGTGCGGACGCTATC ATTGTACACGGTTTCCCGGGCGCGGATTCCGTACGTGCTTGCCTGAACGTTGCAGAAGAAATG GGCCGCGAAGTTTTCCTGCTGACCGAAATGAGCCACCCGGGTGCGGAAATGTTCATCCAGGGC GCGGCTGACGAAATCGCTCGTATGGGTGTAGATCTGGGCGTTAAAAACTACGTGGGTCCGAGC ACCCGTCCGGAACGTCTGTCTCGTCTGCGTGAAATCATTGGCCAGGACTCTTTCCTGATCTCC CCGGGTGTTGGTGCACAGGGTGGCGACCCGGGTGAAACTCTGCGTTTCGCAGATGCGATCATT GTGGGCCGCTCCATCTATCTGGCGGACAACCCGGCTGCTGCTGCGGCTGGCATCATCGAATCT ATCAAAGACCTGCTGAACCCGTAA

Table 13. MtOMPDC H128N protein sequence

>MtOMPDC_H128N
MNRLILAMDLMNRDDALRVTGEVREYIDTVKIGYPLVLSEGMDIIAEFRKRFGCRI IADFKVADIPETNEKICRATFKAGADAIIVHGFPGADSVRACLNVAEEMGREVFLL TEMSNPGAEMFIQGAADEIARMGVDLGVKNYVGPSTRPERLSRLREIIGQDSFLIS PGVGAQGGDPGETLRFADAIIVGRSIYLADNPAAAAAGIIESIKDLLNP

Table 14. Primers used for site-directed mutagenesis of MtOMPDC

H128N forward	CCTGCTGACCGAAATGAGCAACCCGGGTGCGGAAATGTTCATC
H128N reverse	GATGAACATTTCCGCACCCGGGTTGCTCATTTCGGTCAGCAGG
S127Amber/H128N forward	TTTCCTGCTGACCGAAATGTAGAACCCGGGTGCGGAAATGTTC
S127Amber/H128N reverse	GAACATTTCCGCACCCGGGTTCTACATTTCGGTCAGCAGGAAA
T7 forward	TAATACGACTCACTATAGGG
T7 reverse	CTAGTTATTGCTCAGCGGTG

Table 15. Sequence of amber suppressor tRNA and Flexizyme

Amber tRNA	GGCUCUGUAGUUCAGUCGGUAGAACGGCGGACUCUAAAUCC GUAUGUCACUGGUUCGAGUCCAGUCAGAGCCGCCA
dFx	GGAUCGAAAGAUUUCCGCAUCCCCGAAAGGGUACAUGGCGUU AGGU

Table 16. Data collection and refinement statistics for BMP complex of MtOMPDC H128N

Data collection	
Space group	$P2_1$
No. of molecules in asym. unit	2
Cell dimensions	
a (Å)	59.96
b (Å)	63.70
c (Å)	61.74
β(°)	115.30
Resolution (Å)	1.32
No. of unique	98758
reflections	
R_{merge}	0.057
Completeness (%)	99.8
Refinement	
Resolution (Å)	25.0-1.32
R_{cryst}	0.161
R_{free}	0.178
No. atoms	
Protein	3518
Waters	521
Bound ligands	BMP, EDO, CL, MG
Ligands atoms	54
R.m.s deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.08
PDB entry	4LC8

Supplemental Figures

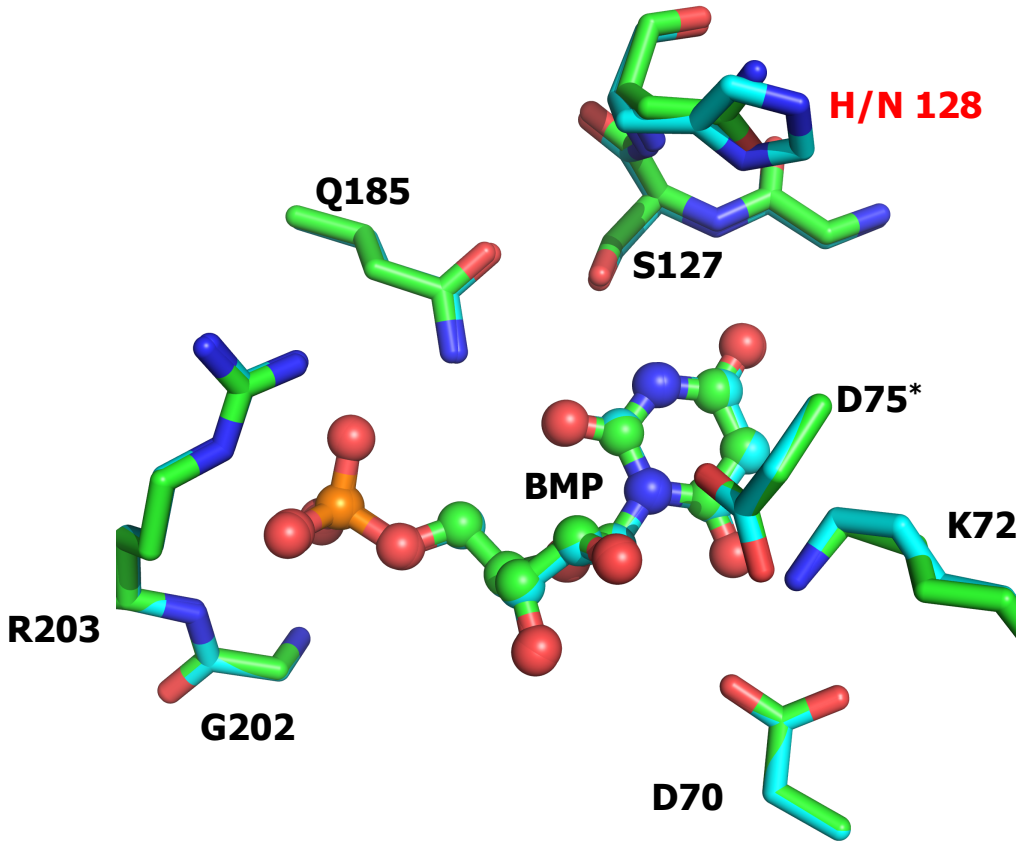


Figure 27. Superposition of active-site of MtOMPDC Wt•BMP (cyan) and MtOMPDC H128N•BMP (green). The PDB coordinates used for this figure were 3LTP (wt) and 4LC8 (H128N). The figure was generated using PyMOL Molecular Graphics System (Schrödinger, LLC).

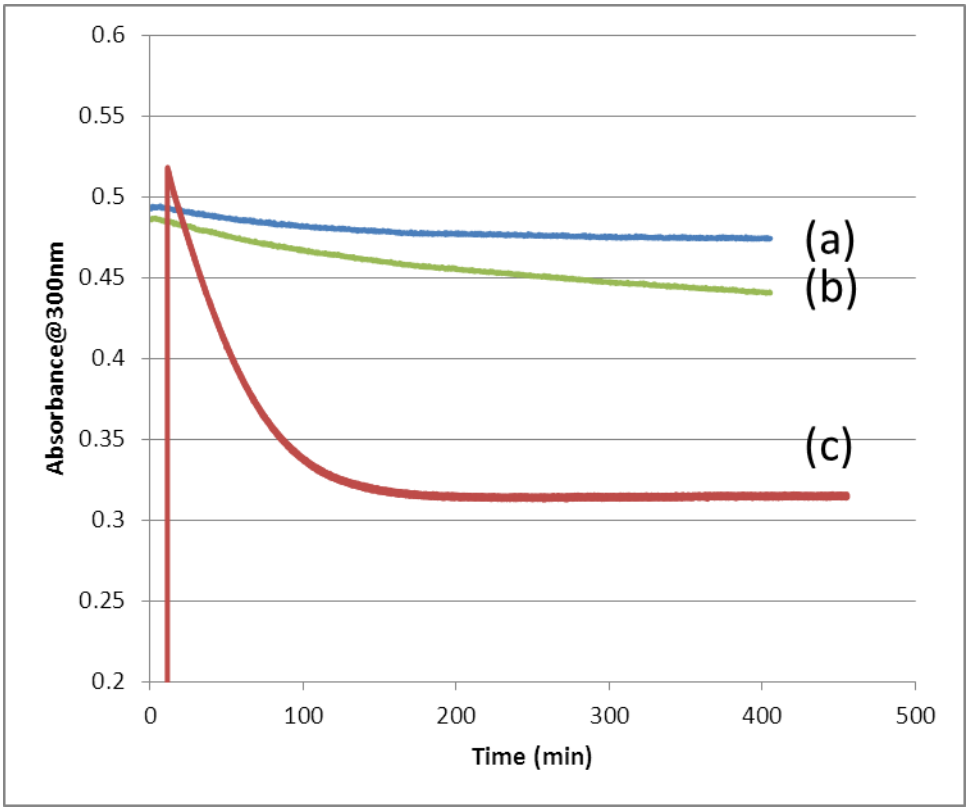


Figure 28. Decay in UV absorbance measured at 300 nm, 25°C, pH 7.1, containing 455 µM starting concentration of FOMP and 10 µL *in vitro* protein synthesis reaction carried out using **a)** No template DNA (background); **b)** S127Amber/H128N template DNA in absence of amber tRNA (negative control); and **c)** S127Amber/H128N template DNA in presence of amber tRNA charged with L-glycerate (^{HO}S).

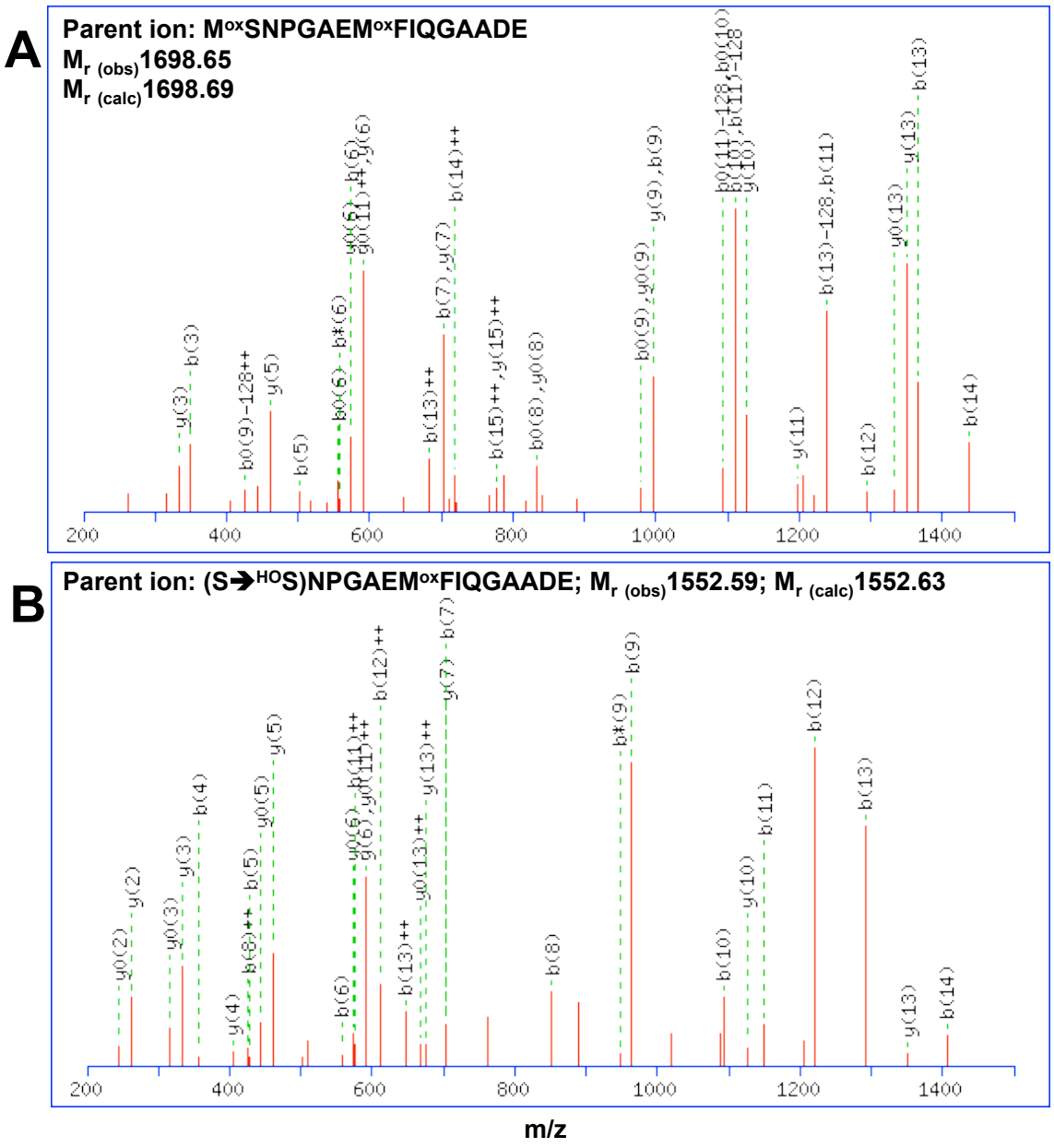


Figure 29. Tandem MS/MS analysis of *in vitro* synthesized wild type (H128N) and S127^{HO}S proteins. Panel A, MS/MS fragmentation of parent ion from wild type protein, corresponding to the sequence M^{ox}SNPGAEM^{ox}FIQGAADE (residues 126-141). Panel B, MS/MS fragmentation of parent ion from S127^{HO}S protein, corresponding to the sequence (S→^{HO}S)NPGAEM^{ox}FIQGAADE (residues 127-141). The observed and calculated monoisotopic masses (M$_r$) of the parent ions are indicated on each spectrum. Assigned b and y fragment ions are numbered according to the modified Roepstorff & Fohlman nomenclature (29, 30).

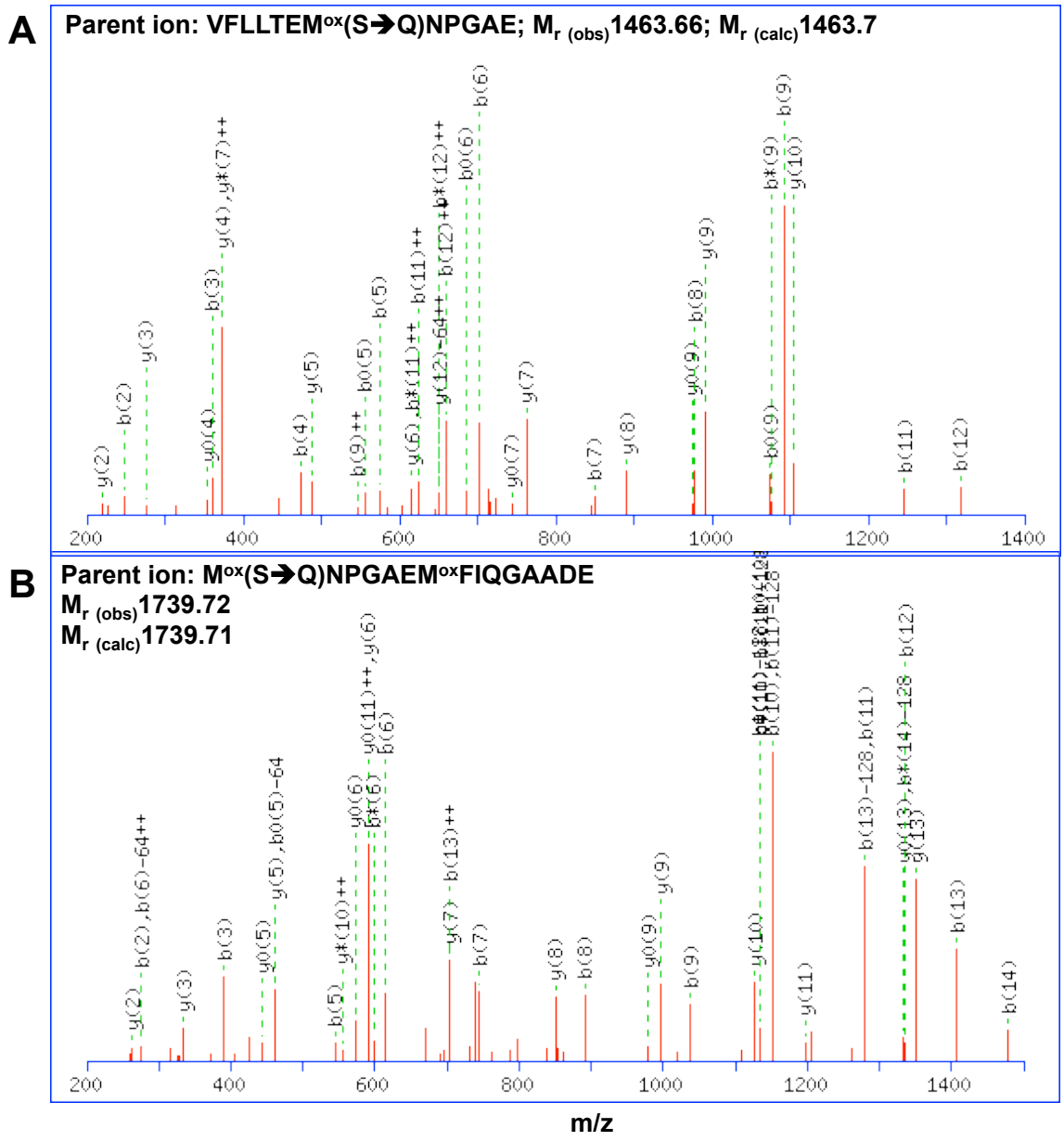


Figure 30. Tandem MS/MS analysis of *in vitro* synthesized protein produced using MtOMPDC S127Amber/H128N DNA, in absence of amber suppressor tRNA. Panel A, MS/MS fragmentation of parent ion, corresponding to the sequence VFLLTEM^{ox}(S→Q)NPGAE. Panel B, MS/MS fragmentation of parent ion, corresponding to the sequence M^{ox}(S→Q)NPGAEM^{ox}FIQGAADE. The observed and calculated monoisotopic masses (M_r) of the parent ions are indicated on each spectrum. Assigned b and y fragment ions are numbered according to the modified Roepstorff & Fohlman nomenclature (29, 30).

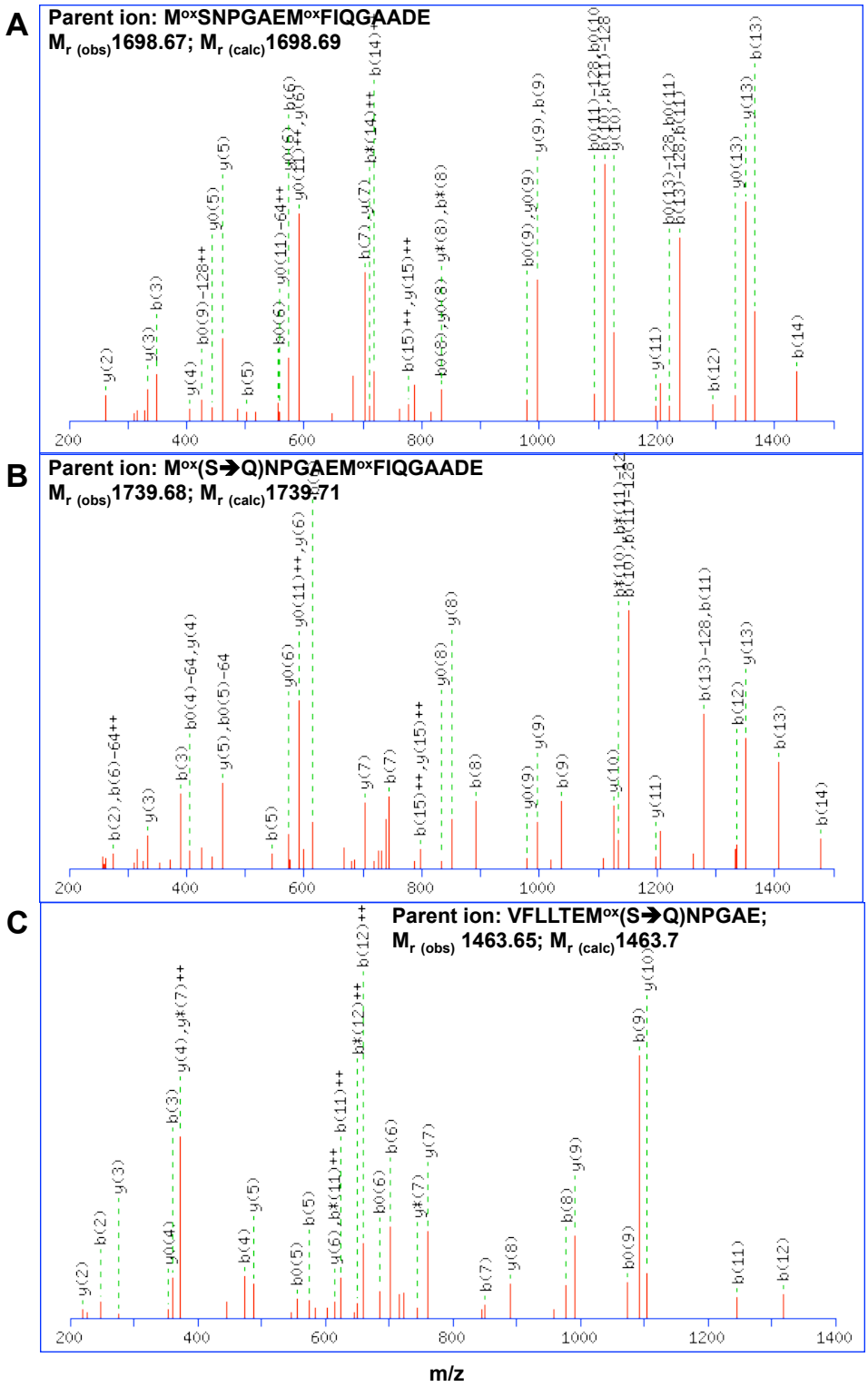


Figure 31. Tandem MS/MS analysis of *in vitro* synthesized protein produced using MtOMPDC S127Amber/H128N DNA and amber suppressor tRNA charged with Ser. Panel A, MS/MS fragmentation of parent ion, corresponding to the sequence

M^{ox}SNPGAEM^{ox}FIQGAADE. Panel B, MS/MS fragmentation of parent ion, corresponding to the sequence M^{ox}(S→Q)NPGAEM^{ox}FIQGAADE. Panel C, MS/MS fragmentation of parent ion, corresponding to the sequence VFLLTEM^{ox}(S→Q)NPGAE. The observed and calculated monoisotopic masses (M_r) of the parent ions are indicated on each spectrum. Assigned b and y fragment ions are numbered according to the modified Roepstorff & Fohlman nomenclature(29, 30).

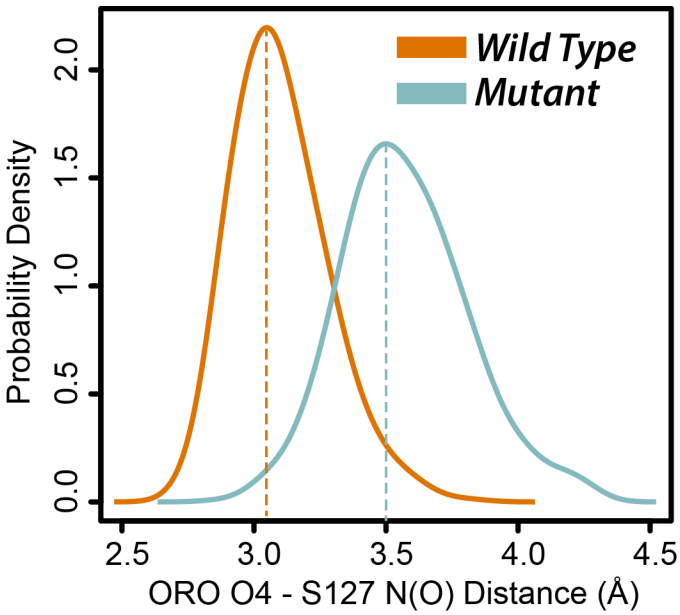


Figure 32. Distribution of the bond distances between O4 and Ser127 N, or ^{HO}S127 O.

DNA	A	B	-	B	B
tRNA	-	S	-	^{HO}S	-

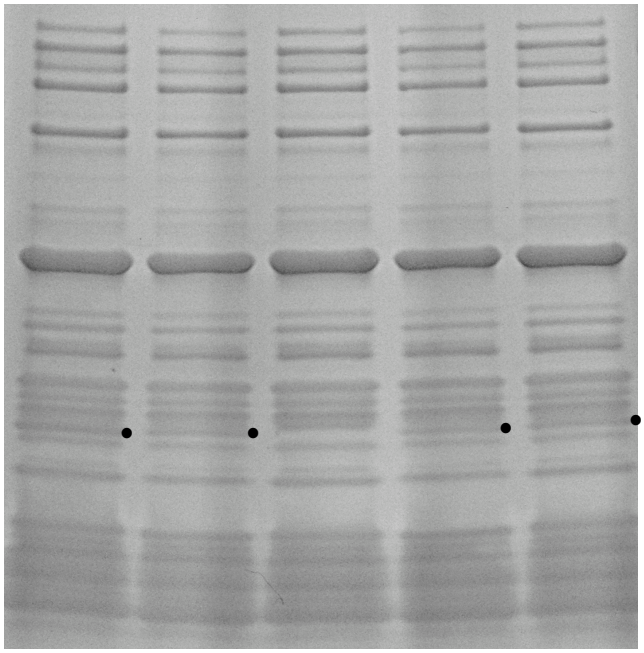


Figure 33. SDS PAGE purification of *in vitro* synthesized proteins for mass spectrometric analysis. A = Wt (H128N); B = S127Amber/H128N. The bands that correspond to the full-length protein used for mass spectrometric analysis are marked by (•).

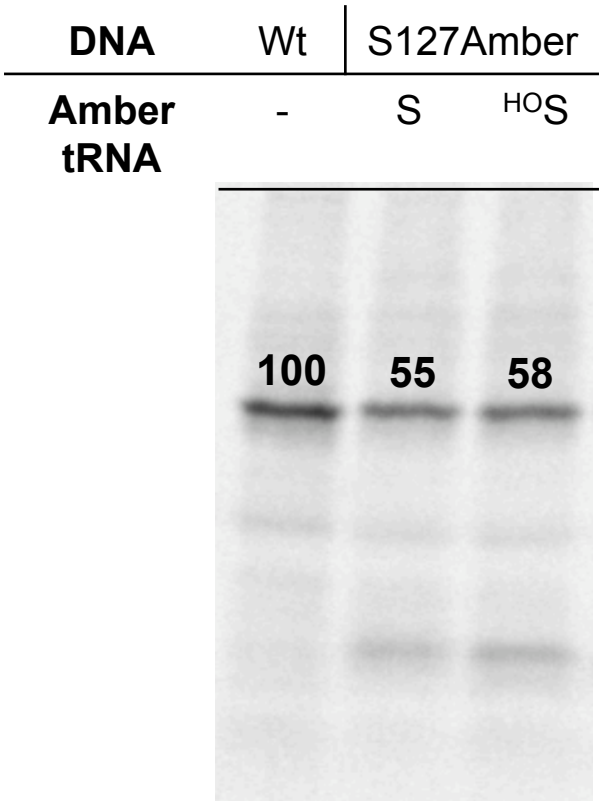


Figure 34. Radioactive image of SDS-PAGE experiments performed on *in vitro* protein synthesis reactions containing [^{35}S]-methionine. Relative expression levels of the full-length protein from reactions containing tRNA charged with L-glycerate (^{HO}S) and Ser (S) compared to wild type (H128N) are indicated above the respective bands.

References

(1) Wood BM, Chan KK, Amyes TL, Richard JP & Gerlt JA (2009) Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: Effect of solvent viscosity on kinetic constants. *Biochemistry* 48(24): 5510-5517.

(2) Wood BM, *et al* (2010) Conformational changes in orotidine 5'-monophosphate decarboxylase: "Remote" residues that stabilize the active conformation. *Biochemistry* 49(17): 3514-3516.

(3) Desai BJ, *et al* (2012) Conformational changes in orotidine 5'-monophosphate decarboxylase: A structure-based explanation for how the 5'-phosphate group activates the enzyme. *Biochemistry* 51(43): 8665-8678.

(4) Goto Y, Katoh T & Suga H (2011) Flexizymes for genetic code reprogramming. *Nat Protoc* 6(6): 779-790.

(5) Ohta A, Murakami H, Higashimura E & Suga H (2007) Synthesis of polyester by means of genetic code reprogramming. *Chem Biol* 14(12): 1315-1322.

(6) Murakami H, Ohta A, Ashigai H & Suga H (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat Methods* 3(5): 357-359.

(7) Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data. *Meth Enzymol* 276: 307-326.

(8) Long F, Vagin AA, Young P & Murshudov GN (2008) BALBES: A molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1): 125-132.

(9) Emsley P & Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1): 2126-2132.

(10) Adams PD*, et al* (2010) PHENIX: A comprehensive python-based system for macromolecular structure solution*. Acta Crystallogr D Biol Crystallogr* 66(Pt 2): 213-221.

(11) Lamzin VS & Wilson KS (1993) Automated refinement of protein models*. Acta Crystallogr D Biol Crystallogr* 49(Pt 1): 129-147.

(12) Hooft RW, Vriend G, Sander C & Abola EE (1996) Errors in protein structures*. Nature* 381(6580): 272.

(13) Chen VB*, et al* (2010) MolProbity: All-atom structure validation for macromolecular crystallography*. Acta Crystallogr D Biol Crystallogr* 66(Pt 1): 12-21.

(14) Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography*. Acta Crystallogr D Biol Crystallogr* 50(Pt 5): 760-763.

(15) Delano WL (2002) The PyMOL molecular graphics system

(16) Wu N & Pai EF (2002) Crystal structures of inhibitor complexes reveal an alternate binding mode in orotidine-5'-monophosphate decarboxylase*. J Biol Chem* 277(31): 28080-28087.

(17) Gao J, Amara P, Alhambra C & Field MJ (1998) A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations*. J Phys Chem A* 102(24): 4714-4721.

(18) MacKerell AD*, et al* (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins*. J Phys Chem B* 102(18): 3586-3616.

(19) Wu N, Mo Y, Gao J & Pai EF (2000) Electrostatic stress in catalysis: Structure and mechanism of the enzyme orotidine monophosphate decarboxylase. *Proc Natl Acad Sci U S A* 97(5): 2017-2022.

(20) Nose S (1983) Constant pressure molecular dynamics for molecular systems. *Mol Phys* 50(5): 1055.

(21) Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev , A* 31(3): 1695-1697.

(22) Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* 72(4): 2384-2393.

(23) Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79(2): 926-935.

(24) Darden T, York D & Pedersen L (1993) Particle mesh ewald: An n·log(N) method for ewald sums in large systems. *J Chem Phys* 98(12): 10089-10092.

(25) Torrie GM & Valleau JP (1977) Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* 23(2): 187-199.

(26) Kirkwood JG (1935) Statistical mechanics of fluid mixtures. *J Chem Phys* 3(5): 300-313.

(27) Wu N, Mo Y, Gao J & Pai EF (2000) Electrostatic stress in catalysis: Structure and mechanism of the enzyme orotidine monophosphate decarboxylase. *Proc Natl Acad Sci U S A* 97(5): 2017-2022.

(28) Kumar S, Rosenberg JM, Bouzida D, Swendsen RH & Kollman PA (1992) THE weighted histogram analysis method for free-energy calculations on biomolecules. I. the method. *Journal of Computational Chemistry* 13(8): 1011-1021.

(29) Roepstorff P & Fohlman J (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* 11(11): 601.

(30) Biemann K (1990) Nomenclature for peptide fragment lons (postive lons). *Methods in Enzymol.* 193: 886.

APPENDIX B**:** SUPPORTING INFORMATION FOR CHAPTER 3

CONFORMATIONAL CHANGES IN OROTIDINE 5'-MONOPHOSPHATE DECARBOXYLASE: A STRUCTURE-BASED EXPLANATION FOR HOW THE 5'-PHOSPHATE GROUP ACTIVATES THE ENZYME

Bijoy J. Desai[a,c], B. McKay Wood[a,c], Alexander A. Fedorov[d], Elena V. Fedorov[d], Bogdana Goryanova[e], Tina L. Amyes[e], John P. Richard[e], Stephen C. Almo[d] and John A. Gerlt[a,b,c]

[a,c]Department of Biochemistry and Institute of Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois-61801.
[b]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.
[d]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461
[e]Department of Chemistry, University of Buffalo, Buffalo, New York 14260

TABLE OF CONTENTS

Supplemental Methods

Assay for Decarboxylation of EO. EO decarboxylation assays were performed by the same method as previously published (2). Before performing the assays the enzyme was dialyzed exhaustively against 50mM MOPS at pH 7.0 and I=0.179(NaCl). Stock concentration of enzyme was determined spectro-photometrically at 280 nm using the proteins molar extinction coefficient (5960 $M^{-1}cm^{-1}$ for Wt and all mutants except ones containing Y206F mutation; or 4470 for Y206F containing mutants). The stock concentration of EO was determined spectro-photometrically at 267 nm using molar extinction coefficient of 9570 $M^{-1}cm^{-1}$. Stock solutions of enzyme and EO were mixed along with buffer (50mM MOPS, pH 7.0; I=0.179) and 0.28 M NaCl to give final reaction condition of 25mM MOPS, pH 7.0; I=0.14; final enzyme concentration ranging from 50 to 500 μM (depending on the mutant) so that the first 5% of the reaction can be observed in a reasonable time; and two EO concentrations of 5mM and 10 mM respectively. The reaction was carried out at 25° C maintained in a water bath. Aliquots from the reaction were taken at regular time intervals. The reaction was quenched by diluting the aliquots 10 fold using quenching buffer (2.5 mM formic acid containing 40 μM 5-fluoro-uracil as internal standard), to give a final pH of 3.8. To remove the enzyme; the quenched reaction aliquots were passed through Millipore Biomax 10 kDa NMWL, 0.5 mL centrifugal filters pre-rinsed with water. The reaction aliquots were analyzed for product formation (1-(β-D-erythrofuranosyl)uracil, EU) by the Beckman Gold System HPLC using a 3 μm (3.9 mm x 150 mm) dC18 column (Waters Atlantis) and 10 mM ammonium acetate (Fluka), pH 4.2, as the isocratic eluent at a flow rate of

1 mL/min. 45 µl of the quenched reaction aliquots were injected and the product observed by setting the HPLC UV detector at 262 nM. The EU concentrations formed at various times were determined by interpolating the area of the EU peak observed, onto a calibration curve generated using uridine (product analog). The calibration curve was generated by injecting various dilutions of uridine (5µM to 100µM) in quenching buffer on the column and measuring the area of the uridine peak. The exact concentration of uridine stock was determined spectro-photometrically at 262 nM using a molar extinction coefficient of 10100 $M^{-1}cm^{-1}$ in 0.1 N HCl. The EU concentration was corrected for dilution due to filtration by comparing the area of the internal standard of each injection to that of unfiltered "blank". From the amount of EU formed at different time intervals, an initial velocity was determined (M s^{-1}). Second order rate constant (k_{cat}/K_M) was determined by dividing the initial velocity by the concentration of enzyme and substrate used in the reaction. The enzyme was ensured to be active throughout the time of the assay by performing OMP decarboxylation assays at the beginning and end of the assays.

Assay for Decarboxylation of FEO. FEO decarboxylation assay of MtOMPDC Wt and single mutants were performed by observing the first order decay of FEO spectro-photometrically at 290 nm in a similar way as done for OMP decarboxylation. Stock concentration of FEO was determined spectro-photometrically at 270 nm using a molar extinction coefficient of 9895 $M^{-1}cm^{-1}$ in 0.1 N HCl. The first order decay curves were fit to obtain the first order rate constant (V_{max}/K_M.) The second order rates (k_{cat}/K_M) was

determined by dividing the first order rate constant with the concentration of enzyme used in the assay.

FEO decarboxylation assay of multiple mutants were performed using the HPLC based assays similar to the one described for EO (2) with slight modifications. The reaction conditions and method used to determine product were the same as EO assay, except a) the HPLC UV detector was set to detect at 270 nm and b) FEU concentrations were determined by interpolating the area of FEU peak observed to the calibration curve generated using 5-fluoro-uridine (product analog). The 5-fluoro-uridine stock concentration was determined spectro-photometrically at 269 nM using a molar extinction coefficient of 9660 $M^{-1}cm^{-1}$ in 0.1 N HCl.

Supplemental Tables

Table 17. Gene sequence of MtOMPDC wild type protein

MtOMPDC Wt gene
ATGAGATCCCGGAGAGTTGATGTTATGGACGTTATGAACAGACTTATACTTGCAATGGAC CTCATGAACCGGGATGATGCCCTCAGGGTCACCGGAGAGGTCAGGGAATACATAGACACG GTCAAGATAGGGTACCCCCTTGTACTCTCAGAGGGTATGGATATCATTGCTGAGTTCAGA AAGAGATTTGGCTGCAGAATCATAGCCGACTTCAAGGTTGCAGATATACCCGAGACCAAT GAAAAGATATGCCGGGCCACCTTCAAGGCGGGGGCTGATGCCATAATAGTCCACGGATTC CCGGGAGCAGACAGTGTCAGGGCCTGCCTCAATGTCGCAGAGGAAATGGGACGTGAGGTC TTCCTCCTGACAGAGATGTCACACCCGGGGGCAGAGATGTTCATACAGGGCGCTGCAGAT GAAATAGCCAGAATGGGGGTCGATCTTGGTGTCAAAAATTATGTTGGCCCATCCACAAGA CCTGAAAGGCTTTCAAGGCTGAGGGAAATCATAGGTCAGGATTCATTTCTCATATCCCCC GGTGTGGGAGCCCAGGGAGGAGACCCAGGGGAGACCCTCAGGTTCGCAGATGCCATAATA GTTGGAAGATCCATCTACCTTGCAGATAACCCTGCAGCTGCAGCAGCAGGGATAATAGAA TCCATTAAAGACCTTCTGAATCCCTGA

Table 18. Protein sequence of MtOMPDC

MtOMPDC Wt
MRSRRVDVMDVMNRLILAMDLMNRDDALRVTGEVREYIDTVKIGYPLVLSEGMDIIAEFR KRFGCRIIADFKVADIPETNEKICRATFKAGADAIIVHGFPGADSVRACLNVAEEMGREV FLLTEMSHPGAEMFIQGAADEIARMGVDLGVKNYVGPSTRPERLSRLREIIGQDSFLISP GVGAQGGDPGETLRFADAIIVGRSIYLADNPAAAAAGIIESIKDLLNP

APPENDIX C: SUPPORTING INFORMATION FOR CHAPTER 4

COMPUTATIONAL, STRUCTURAL AND KINETIC EVIDENCE THAT *VIBRIO VULNIFICUS* FRSA IS NOT A COFACTOR-INDEPENDENT PYRUVATE DECARBOXYLASE

Whitney F. Kellett[a], Elizabeth Brunk[b], Bijoy J. Desai[c], Alexander A. Fedorov[d], John A. Gerlt[c], Steven C. Almo[d], Ursula Rothlisberger[b], and Nigel G. J. Richards[a,1]

[a]Department of Chemistry, University of Florida, Gainesville, FL 32611, USA, Laboratory of Computational Chemistry and Biochemistry

[b]Ecole Polytechnique Fédérale Lausanne, CH-1015 Lausanne, Switzerland,

[c]Departments of Chemistry & Biochemistry, Institute for Genomic Biology, University of Illinois at Urbana-Champaign ,

[d]Albert Einstein College of Medicine,Bronx, NY 10461, USA.

[1]Present address: Department of Chemistry & Chemical Biology, 402. N. Blackford. St., Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA

Table of Contents

Supplemental Methods

Obtaining an initial model of the FrsA/pyruvate complex

Coordinates for the free enzyme monomer were obtained from the PDB (entry: 3MVE) and all crystallographic waters and counter-ions were removed prior to the addition of hydrogen atoms using the tLEAP module of the AMBER11 software package (1, 2). The resulting structure was placed in a box of explicit TIP3P water molecules (3) with two chloride ions to yield a system with zero net charge. After constrained energy minimization, the FrsA structure was carefully equilibrated using a series of constrained molecular dynamics simulations (NPT ensemble). These calculations all employed the AMBER99 force field (4, 5), as implemented in the Sander module of the AMBER11 software package (1), and the Particle-Mesh Ewald algorithm to treat electrostatic interactions (6). Force field parameters for pyruvate were generated using the Generalized AMBER Force Field (GAFF) (7), and then the GLIDE tool (8) in the MAESTRO package was used to generate models of the FrsA/pyruvate complex. The docking calculations employed an active site consisting of the residues (Arg-53, Asp-203, Arg-272, Phe-273, Phe-303, Tyr-316, Leu-320, Met-338 and Trp-341) proposed in the original description of FrsA structure and function (9). Subsequent energy minimization of eight FrsA/pyruvate complexes followed by constrained MD equilibration (over a period of 40 ns in the NPT ensemble) gave the same final structure, which was then used as the starting point for QM/MM simulations (**Figure 35**). Coordinates for this FrsA/pyruvate complex in PDB format (10) are available from the authors on request.

CPMD simulations of pyruvate decarboxylation in the FrsA/pyruvate complex

The QM/MM simulations employed an extension of CPMD (12, 13). Atoms in the QM region, comprising pyruvate, three active site waters, and the Tyr-316 side chain capped at C_α with a monovalent pseudopotential (13), were described by the BLYP functional (14, 15) and norm-conserving Martins-Trouiller pseudopotentials (16) with dispersion-corrected atom-centered potentials (17-20). Wavefunctions were expanded in a plane wave basis set with a 70 Ry cutoff inside an orthorhombic box (37 Å x 22 Å x 22 Å), and long-range interactions were decoupled using the Martyna-Tuckerman scheme (21). The remaining protein atoms and explicit solvent water molecules were described by a classical AMBER99 force field (4, 5). QM/MM simulations were performed at constant pressure and temperature, using a Nosé Hoover thermostat, after slow heating to 300K over 2-3 ps. In these heating simulations, however, the Tyr-316 hydroxyl spontaneously transferred a proton to the pyruvate carboxylate when the temperature exceeded 250 K. Control calculations, using a constraint on the O-H bond of the Tyr-316 moiety, demonstrated that the proton transfer event was not due to the starting configuration. Thus, when the constraint on the O-H bond was released at 300K, spontaneous proton transfer again took place. The resulting system containing neutral pyruvic acid was then equilibrated for 2 ps before constrained MD simulations were performed for thermodynamic integration (22, 23) using a reaction coordinate chosen to be the bond distance between C1 and C2 of the pyruvate molecule. Hence, constraints were employed at distances spanning 1.55 to 4.24 Å (in increments of 25

pm) with sampling at each constrained distance for a period of 1 ps. The free energy profile was obtained by integrating the constraint forces over the respective distances.

CPMD simulations of pyruvate decarboxylation in water

Pyruvate monoanion was placed in a box of explicit TIP3P water molecules (3) with a single potassium ion to yield a system with zero net charge, and the resulting system equilibrated using a series of MD simulations (over a period of 40 ns in the NPT ensemble). These calculations all employed the AMBER99 force field (4, 5), as implemented in the Sander module of the AMBER11 software package (1), and the Particle-Mesh Ewald algorithm to treat electrostatic interactions (6). GAFF (7) parameters were employed to describe pyruvate in these equilibration studies. The equilibrated system was then employed in QM/MM simulations (12, 13). Atoms in the QM region comprised the pyruvate monoanion and a water molecule, which was hydrogen bonded to the carboxylate group of pyruvate. These QM atoms were described by the BLYP functional (14, 15) and norm-conserving Martins-Trouiller pseudopotentials (16) with dispersion-corrected atom-centered potentials (17-20). Wavefunctions were expanded in a plane wave basis set with a 70 Ry cutoff inside an orthorhombic box (12 Å x 7 Å x 7 Å), and long-range interactions were decoupled using the Martyna-Tuckerman scheme (21). The remaining water molecules were described by a classical AMBER99 force field (4, 5). QM/MM simulations were performed at constant pressure and temperature, using a Nosé Hoover thermostat, after heating to 300K. After equilibration, constrained MD simulations were performed for

thermodynamic integration (22, 23) using a reaction coordinate chosen to be the bond distance between C1 and C2 of the pyruvate molecule. Hence, constraints were employed at distances spanning 1.6 to 4.5 Å (in ten increments) with sampling at each constrained distance for a period of 1 ps. The free energy profile was obtained by integrating the constraint forces over the respective distances

pKa estimates using Poisson-Boltzmann calculations

The protonation state of all protein residues in the X-ray crystal structure of unliganded FrsA (3MVE), the equilibrated model of the FrsA/pyruvate complex, and several different "frames" from classical MD trajectories for unliganded FrsA and the FrsA/pyruvate complex were calculated (**Table 19**) using two well-established methods, H++ (24, 25) and Adaptive Poisson-Boltzmann (APB) (26). These calculations suggest that the pK_a of the Tyr-316 side chain is lowered by approximately one unit when pyruvate is bound to the protein (**Table 19**). Superimposing different MD snapshots of the active site is consistent with the view that electrostatic repulsion between the substrate carboxylate and the Asp-203 side chain, and steric interactions with Met-312, positions pyruvate close to Tyr-316 and Arg-273 (**Figure 37**).

Protein expression and purification

Method A

The gene encoding the FrsA from *Vibrio vulnificus* in the plasmid pQE-FrsA was generously provided by Dr. Sun-Shin Cha (Marine Biotechnology Research Center, Korean Ocean Research & Development Institute) and expressed in *Escherichia coli* JM109 cells containing the F' episomal element, as reported previously (9). F' is essential for protein expression, and JM109 cells containing this element were identified by screening on minimal growth media containing proline. After transformation with pQE-FrsA, starter cultures of JM109 cells were grown overnight in M9 minimal media supplemented with thiamin (5 µg/mL), L-proline (5 µg/mL) and ampicillin (100 µg/mL). These starter cultures were used to inoculate 3 L of M9 minimal media supplemented with thiamin, L-proline and ampicillin (100 µg/mL). The cultures were grown with aeration at 37 °C to mid-log phase (O.D of 0.6) and then heat shocked at 42 °C for 15 min before the addition of IPTG (0.5 mM) to induce protein expression. Cells were harvested after 6 h, lysed using sonication, and subsequently centrifuged at 5000x *g* to obtain a clear lysate that was added to Ni-NTA resin suspended in a phosphate wash buffer (50 mM NaH_2PO_4 containing 10 mM imidazole and 500 mM KCl, pH 7). After 1 h incubation on ice, FrsA was eluted from the resin (50 mM NaH_2PO_4 containing 250 mM imidazole and 500 mM KCl, pH 7) and protein-containing fractions were pooled and dialyzed into 50 mM NaH_2PO_4 containing 500 mM KCl and 1 mM DTT, pH 7. The inclusion of DTT in the elution buffer was to prevent formation of the FrsA dimer by a

disulfide bond between the cysteine residues, which had been reported to be inactive (9).

Method B

The codon-optimized gene encoding the FrsA from *Vibrio vulnificus* was synthesized (GenScript, Piscataway NJ) and ligated into the PUC57 plasmid with NdeI and BamHI restriction sites at the 5' and 3' ends, respectively. The resulting plasmid was subjected to digestion with NdeI and BamHI and ligated into the pET15b expression vector (N-terminal His_6-tag). The ligated plasmid was transformed into *Escherichia coli* XL1 blue cells, and a single colony was inoculated into 5 mL of LB containing ampicillin (100µg/mL) and allowed to grow overnight at 37 °C. The plasmid encoding FrsA was then transformed into *Escherichia coli* BL21 (DE3) cells and starter cultures prepared from single colonies. These were used to inoculate 4 L of LB containing ampicillin (100 µg/mL), and resulting cultures were grown at 37 °C with aeration to mid-log phase (O.D. ~ 0.6) before the addition of IPTG (1 mM final concentration). Cells were harvested after 4 h, and lysed by sonication. The clarified lysate was then loaded onto Ni-NTA resin and FrsA eluted using an imidazole gradient. The fractions containing FrsA (determined by SDS PAGE) were pooled, dialyzed against phosphate buffered saline containing 3 mM DTT, and the His-tag cleaved using thrombin (Promega, Madison WI). The resulting preparation of FrsA was dialyzed against storage buffer (20 mM HEPES containing 150 mM NaCl and 3 mM DTT, pH 7.5). After concentration to ~ 10 mg/mL, the purified FrsA was stored in 20 µL aliquots at -80 °C.

Kinetic assays of recombinant, wild type FrsA

Method A: Membrane-Inlet Mass Spectrometry (MIMS) for direct detection of CO_2

The real-time production of $^{13}CO_2$ was monitored by MIMS (27) when 1 mM ^{13}C-labeled pyruvate was incubated with 50 nM recombinant FrsA dissolved in 50 mM NaH_2PO_4 buffer, pH 7, containing 500 mM KCl and 1 mM DTT, over a period of 12 min at 37°C. No detectable decarboxylase activity was observed.

Method B: Spectrophotometric assay for pyruvate decarboxylation. Decarboxylation of pyruvate to acetaldehyde was measured using a coupled assay with yeast alcohol dehydrogenase (Sigma-Aldrich, St. Louis MO) by monitoring the decay in NADH absorbance at 340 nm. Assays were carried out in 50 mM sodium phosphate, pH 7.0; containing 1 mM DTT; 5 mM pyruvate; 0.2 mM NADH; 3.7 units of alcohol dehydrogenase and FrsA ranging from 1 nM to 1 μM. No detectable activity was observed. At the end of each assay, acetaldehyde was added to ensure that the alcohol dehydrogenase was active.

Method C: ^1H NMR detection of the FrsA-catalyzed pyruvate decarboxylation. ^1H NMR spectroscopy was used to detect acetaldehyde formation when 5 μM FrsA was incubated for 90 min at 20 °C with 10 mM pyruvate in D_2O with 50 mM phosphate, pD 7.0, containing 1mM DTT. Although signals corresponding to the methyl group protons

of unhydrated and hydrated pyruvate anion were detected, no resonance(s) were detected for the protons in acetaldehyde.

Crystallization, data collection, structure determination and crystallographic refinement

Crystals of recombinant, wild type FrsA were grown at room temperature using the "sitting drop" method. In these experiments, the protein solution contained FrsA (20 mg/mL), 150 mM NaCl and 3 mM DTT dissolved in 20 mM HEPES buffer, pH 7.5 and the percipient was composed of 20% PEG monomethyl ether 2000 and 200 mM trimethylamine N-oxide in 0.1 M Tris buffer, pH 8.5. Crystals appeared in 3-4 days and exhibited diffraction consistent with the space group $P2_1$, with two polypeptides per asymmetric unit. Prior to data collection, the crystals of FrsA were transferred to cryoprotectant solutions, composed of their mother liquids containing 20% glycerol, and flash-cooled in a nitrogen stream. The X-ray diffraction data set (**Table 20**) was collected at the NSLS X29A beamline (Brookhaven National Laboratory) on the 315Q CCD detector. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK (28).

The FrsA X-ray crystal structure (**Figure 38**) was solved using only input diffraction and sequence data by molecular replacement with the fully automated molecular replacement pipeline BALBES (29). The search model employed by BALBES was that of a previously reported structure for FrsA (3MVE) (9). After obtaining a partially refined

structure from BALBES, several iterative cycles of refinement were then performed. Thus, model rebuilding was performed with COOT (30) and refined with PHENIX (31), followed by automatic model rebuilding using the ARP software package (32). The quality of the final structure was verified with omit maps, and stereochemistry checked with WHATCHECK (33) and MOLPOBITY (34). All loops are well defined in the current FrsA structure, with only the backbone torsion angles of Asn-108 adopting values in the disallowed region of the Ramachandran plot. The program LSQKAB 35) was used for structural superposition, and figures with electron density maps were prepared using PYMOL. Final crystallographic refinement statistics are shown in **Table 20**.

Supplemental Tables

Table 19. pK_a values computed for Tyr-316 in unliganded FrsA (3MVE) and various MD trajectory "snapshots" of the FrsA/pyruvate complex, and for pyruvate in the final model of the FrsA/pyruvate complex.

Residue	Calculated pKa (Method A)[a]	Calculated pKa (Method B)[a]
Tyr-316 (3MVE)	9.6	10.5
Tyr-316 (FrsA/pyruvate: 18.60 ns)	9.8	9.9
Tyr-316 (FrsA/pyruvate: 18.75 ns)	9.9	9.9
Tyr-316 (FrsA/pyruvate: 18.90 ns)	9.0	10.1
Tyr-316 (FrsA/pyruvate: 19.00 ns)	9.2	9.2
Pyruvate (FrsA/pyruvate: 19.00 ns)	n.d.	3.8

[a] H^{++} (http://biophysics.cs.vt.edu/H++)
[b] APBS (ref. 26)

Table 20. Crystallographic data collection and refinement statistics.[a]

Data Collection	
Resolution (Å)	∞-1.95
X-ray Source	NSLS X29A
Wavelength (Å)	1.075
Space Group	$P2_1$
Molecules in asymmetric unit	2
Cell dimensions	a = 42.66 Å, b = 104.54 Å, c = 82.30 Å, β = 91.18°
Unique reflections observed	51363
Completeness (%)	97.8
R_{merge}[b]	0.083
Refinement	
Protein residues/water atoms per asu	6405/274
Other ligands per asu	2
Reflections (work/free)	51363/2608
R_{cryst}/R_{free} (%)	17.9/22.7
Resolution (Å)	25.00-1.95
Average B-factor (Å^2)	29.2
Protein (Å^2)	29.3
Hexanoic acid (Å^2)	32.9
PEG monomethyl ether (Å^2)	47.5
RMSD[c] Bond lengths (Å)	0.008
RMSD Bond angles (°)	1.05

[a] Coordinates are deposited as PDB entry 4I4C.
[b] $R_{merge} = \sum_{hkl} \sum_i | I_{hkl,i} - \langle I_{hkl} \rangle I | / \sum_{hkl} \sum_i | I_{hkl,i} |$, where $\langle I_{hkl} \rangle$ is the mean intensity of the multiple $I_{hkl,i}$ observations for symmetry-related reflections.
[c] RMSD: root mean square deviation.

Supplemental Figures

Figure 35. Ribbon representation of the FrsA/pyruvate model complex showing the location of the bound ligand. Pyruvate is rendered as cylinders colored by atom type. Color scheme: C – cyan; H – white; O – red. The figure was produced with the VMD software package (11).

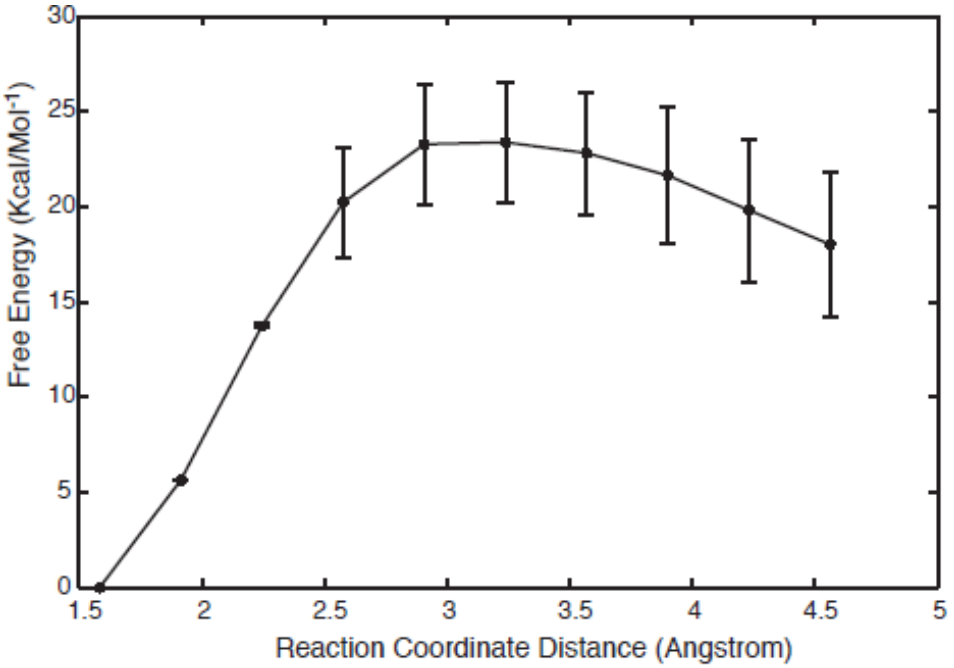


Figure 36. Free energy profile (kcal/mol) for cleavage of the pyruvate C-C bond in water, as computed by thermodynamic integration. Error bars show the statistical sum of errors associated with the calculated free energy.

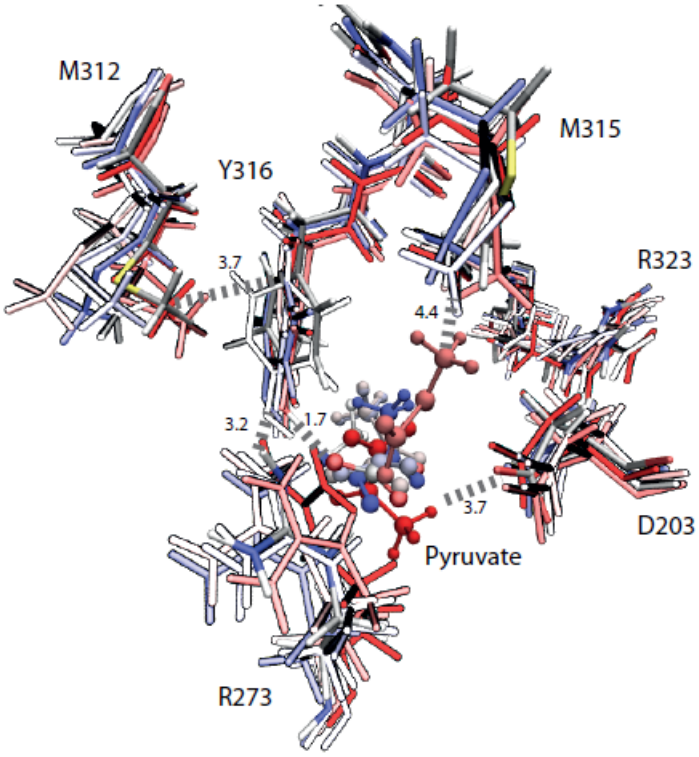


Figure 37. Superimposed snapshots of the active site from the MD trajectory of the equilibrated FrsA/pyruvate complex. The average distance between pyruvate and the Asp-203 side chain is shown by the dashed line. Residue coloring reflects the time of the snapshot, ranging from red (early) to blue (late).

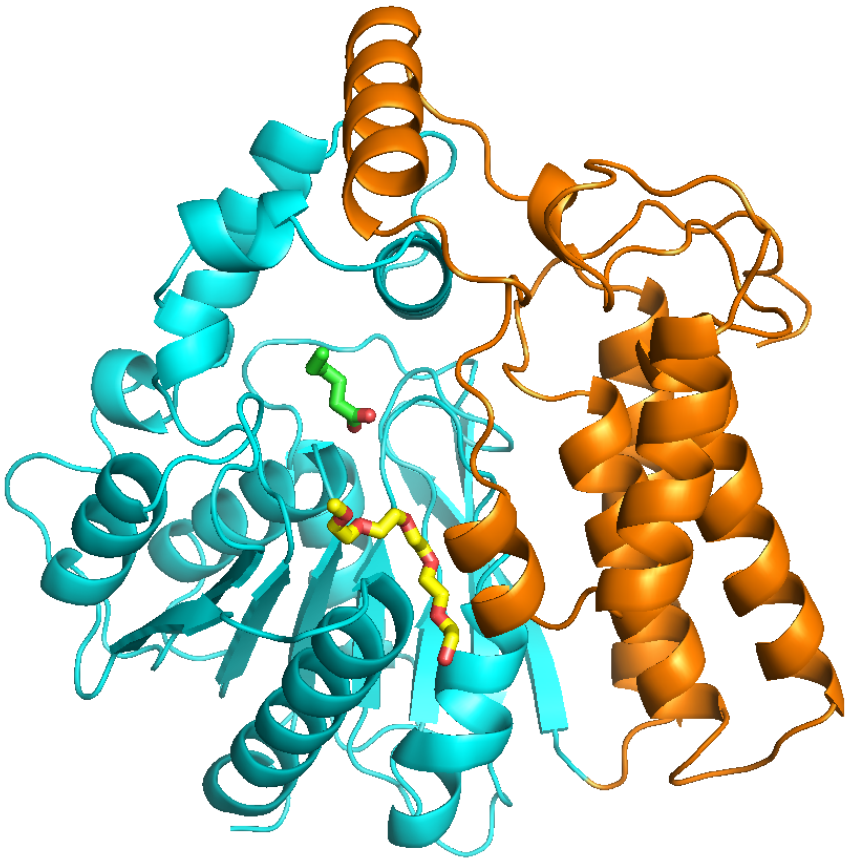


Figure 38. Ribbon representation of the FrsA monomer containing putative ligands. The N- (1-165) and C-terminal (166-415) domains are colored brown and cyan, respectively. Bound PEG monomethyl ether ($H_3C[OCH_2CH_2]_5OH$) and hexanoic acid ($H_3CCH_2CH_2CH_2CH_2$ COOH) ligands are rendered as cylinders colored yellow and green, respectively, except for oxygen atoms, which are shown in red. Bound hexanoate is positioned in a very similar manner to pyruvate in the equilibrated computational model of the FrsA/pyruvate complex Figure 35. The image was produced with PYMOL (DeLano Scientific Software LLC, Palo Alto, CA).

References

(1) Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A. Simmerling, C., Wang, B., and Woods, R. J. (2005) The AMBER biomolecular simulation programs, *J. Comput. Chem. 26*, 1668-1688.

(2) Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Debolt, S., Ferguson, D., Siebel, G., and Kollman, P. A. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, *Computer Phys. Commun. 91*, 1-41.

(3) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926-935.

(4) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple AMBER force fields and development of improved backbone parameters, *Proteins: Struct. Funct. Genet. 65*, 712-725.

(5) Ponder, J. W., and Case, D. A. (2003) Force fields for protein simulations, *Adv. Prot. Chem. 66*, 27-85.

(6) Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald. An NlogN method for ewald sums in large systems, *J. Chem. Phys. 98*, 10089-10092.

(7) Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general AMBER force field, *J. Comput. Chem. 25*, 1157-1174.

(8) Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shaw, D. E., Shelley, M., Perry, J. K., Francis, P., and Shenkin, P. S. (2004) GLIDE: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J. Med. Chem. 47*, 1739-1749.

(9) Lee, K.-J., Jeong, C.-S., An, K. J., Lee, H.-J., park, S.-J., Seok, Y.-J., Kim, P., Lee, J.-H., Lee, K.-H., and Cha, S.-S. (2011) FrsA functions as a cofactor-independent decarboxylase to control metabolic flux, *Nat. Chem. Biol. 7*, 434-436.

(10) Berman, H. M., Kleywegt, G. J., Nakamura, H., and Markley, J. L. (2012) The Protein Databank at 40: Reflecting on the past to prepare for the future, *Structure 20*, 391-396.

(11) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD – Visual Molecular Dynamics, *J. Molec. Graph. 14*, 33-38.

(12) Car, R., and Parrinello, M. (1985) Unified approach for molecular dynamics and density functional theory, *Phys. Rev. Lett. 55*, 2471-2474.

(13) Laio, A., VandeVondele, J., and Rothlisberger, U. (2002) A Hamiltonian coupling scheme for hybrid Car-Parrinello molecular dynamics simulations, *J. Chem. Phys. 116*, 6941–6948.

(14) Becke, A. D. (2008) Density functional exchange energy approximation with correct asymptotic behavior, *Phys. Rev. A 38*, 3098-3100.

(15) Lee, C., Yang, W., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation energy formula into a functional of the electron density, *Phys. Rev. B 37*, 785-789.

(16) Trouiller, N., and Martins, J. L. (1991) Efficient pseudopotentials for plane-wave calculations. 2. Operators for fast iterative diagonalization, *Phys. Rev. B. 43*, 8861–8869.

(17) Lin, I.-C., Seitsonen, A. P., Coutinho-Neto, M. D., Tavernelli, I, and Rothlisberger, U. (2009) Importance of van der Waals interactions in liquid water, *J. Phys. Chem. B 113*, 1127-1131.

(18) von Lilienfeld, O. A., Tavernelli, I., Rothlisberger, U., and Sebastiani, D. (2007) Optimization of effective atom centered potentials for London dispersion forces in density functional theory, *Phys. Rev. Lett. 93*, 153004.

(19) Lin, I.-C., Coutinho-Neto, M. D., Felsenheimer, C., von Lilienfeld, O. A., Tavernelli, I., and Rothlisberger, U. (2007) Library of dispersion-corrected atom-

centered potentials for generalized gradient approximation functionals: Elements H, C, N, O, He, Ne, Ar, and Kr, *Phys. Rev. B. 75*, 205131.

(20) Arey, J. S., Aeberhard, P. C., Lin, I.-C., and Rothlisberger, U. (2009) hydrogen bonding described using dispersion-corrected density functional theory, *J. Phys. Chem. B 113*, 4726-4732.

(21) Martyna, G. J., and Tuckerman, M. E. (1999) Molecular dynamics algorithms for path integrals at constant pressure, *J. Chem. Phys. 110*, 3275-3290.

(22) Sprik, M., and Ciccotti, G. (1998) Free energy from constrained molecular dynamics, *J. Chem. Phys. 109*, 7737-7744.

(23) Kirkwood, J. G. (1935) Statistical mechanics of fluid mixtures, *J. Chem. Phys. 3*, 300-313.

(24) Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. (2012) H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulation, *Nucl. Acids Res. 40*, W537-W541.

(25) Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules, *Proteins: Struct. Funct. Genet. 63*, 928-938.

(26) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome, *Proc. Natl. Acad. Sci., USA 98*, 10037-10041.

(27) Moral, M. E. G., Tu, C. K., Richards, N. G. J., and Silverman, D. N. (2011) Membrane inlet for mass spectrometric measurement of catalysis by enzymatic decarboxylases, *Analyt. Biochem. 418*, 73-77.

(28) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307-326.

(29) Long, F., Vagin, A., Young, P., and Murshudov, G.N. (2008) BALBES: a molecular replacement pipeline, *Acta Crystallogr. D 64*, 125-132.

(30) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr. D 60,* 2126-2132.

(31) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D 66*, 213-221.

(32) Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models, *Acta Crystallogr. D 49*, 129-147.

(33) Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) Errors in protein structures, *Nature 381*, 272-273.

(34) Chen, V. B., Arrendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D.C. (2010)

MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D 66*, 12-21.

(35) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCCP4 suite and current developments, *Acta Crystallogr. Sect. D 67*, 235-242.